



NOAA HAZARDOUS WEATHER **TESTBED**

EXPERIMENTAL FORECAST PROGRAM **SPRING EXPERIMENT 2010**

http://hwt.nssl.noaa.gov/Spring_2010/

HWT Facility – National Weather Center
17 May - 18 June 2010

Program Overview **and** **Operations Plan**

5/22/10 Version

Steven Weiss¹, Patrick Marsh², Chris Melick¹, Israel Jirak¹, Adam Clark², Chris Siewert¹, Ryan Sobash³, Andy Dean¹, Jack Kain², Gregory Carbin¹, Tara Jensen⁴, John Hart¹, Jason Levit⁵, Bruce Entwistle⁵, David Novak⁶, Mike Bodner⁶, and Faye Barthold⁶

- (1) NOAA/NWS/NCEP Storm Prediction Center, Norman, Oklahoma
- (2) NOAA/OAR National Severe Storms Laboratory, Norman, Oklahoma
- (3) School of Meteorology, University of Oklahoma, Norman, Oklahoma
- (4) NCAR/NOAA Developmental Testbed Center, Boulder, Colorado
- (5) NOAA/NWS/NCEP Aviation Weather Center, Kansas City, Missouri
- (6) NOAA/NWS/NCEP Hydrometeorological Prediction Center, Camp Springs, Maryland

Table of Contents

1. The NOAA Hazardous Weather Testbed	3
2. Historical Perspective	5
3. Spring Experiment Background and Motivation	6
<i>Operational Forecasting of Severe Convective Storms: Current State and Challenges</i>	6
<i>The Need for More Detailed Thunderstorm Forecasts</i>	7
<i>Evaluation of Deterministic Convection-Allowing NWP in the Spring Experiment</i>	8
<i>Hourly Maximum Fields from WRF Models</i>	9
<i>Radar Assimilation into Convection-Allowing Models</i>	10
<i>Convective Predictability on the Grid Scale and Storm Scale Ensemble Forecast System</i>	11
4. Additional HWT Focus Areas in 2010 – Aviation-Impacts and QPF	12
<i>Applications to Thunderstorm Impacts on Warm-Season Aviation Operations</i>	12
<i>Applications to Quantitative Precipitation Forecasting</i>	13
6. Experimental Models.....	17
<i>CAPS Models</i>	17
<i>4 km Storm Scale Ensemble Forecast, 1 km WRF-ARW, and VORTEX2 4 km WRF-ARW runs at 09, 12, 15 and 18 UTC</i>	17
<i>EMC 4 km WRF-NMM Model</i>	21
<i>NSSL 4 km WRF-ARW Model</i>	22
<i>GSD 3 km High Resolution Rapid Refresh (HRRR) Model</i>	24
<i>NCAR 3 km WRF-ARW Model</i>	25
7. GOES-R Proving Ground Products	27
8. Objectives and Goals	28
<i>Severe Convective Storm Component (Leader: SPC)</i>	28
<i>Aviation Thunderstorm-Impacts Component (Leader: AWC)</i>	29
<i>QPF Component (Leader: HPC)</i>	30
9. Spring Experiment Web Site	31
10. Dates and Location of the Spring Experiment.....	31
11. Daily Operations Schedule	31
<i>Severe Convective Storms Component</i>	32
<i>Aviation Impacts Component</i>	33
<i>QPF Component</i>	34
12. Acknowledgments.....	35
Appendix B: Instructions for Creating and Submitting Experimental Severe Thunderstorm Forecasts.....	43
Appendix C: Instructions for Creating and Submitting Experimental Aviation Thunderstorm Forecasts.....	44
Appendix D: Instructions for Creating and Submitting Experimental QPF Products	46
Appendix E: Practically Perfect Forecasts.....	47
Appendix F: WRF Model Identification of Convective Storms with Rotating Updrafts – Computation of Updraft Helicity	50
Appendix G: DTC Verification Metrics Summary.....	52

1. The NOAA Hazardous Weather Testbed

NOAA’s Hazardous Weather Testbed (HWT) is a facility jointly managed by the National Severe Storms Laboratory (NSSL), the Storm Prediction Center (SPC), and the NWS Oklahoma City/Norman Weather Forecast Office (OUN) within the National Weather Center building on the University of Oklahoma South Research Campus. The HWT is designed to accelerate the transition of promising new meteorological insights and technologies into advances in forecasting and warning for hazardous mesoscale weather events throughout the United States. The HWT facilities include a combined forecast and research area situated between the operations rooms of the SPC and OUN, and a nearby development laboratory. The facilities support enhanced collaboration between research scientists and operational weather forecasters on specific topics that are of mutual interest.

The HWT organizational structure is composed of three primary overlapping program areas (Fig. 1). The first program area focuses on application of cutting edge numerical weather prediction models to improve severe weather forecasts under the auspices of the Experimental Forecast Program (EFP), and the second program tests research concepts and technology specifically aimed at short-fused warnings of severe convective weather under auspices of the Experimental Warning Program (EWP). A key NWS strategic goal is to extend warning lead times under the concept of “Warn-on-Forecast” through the development and application of convection-allowing numerical models to extend short-term predictability of hazardous convective weather. This provides a natural overlap between the EFP and EWP activities.

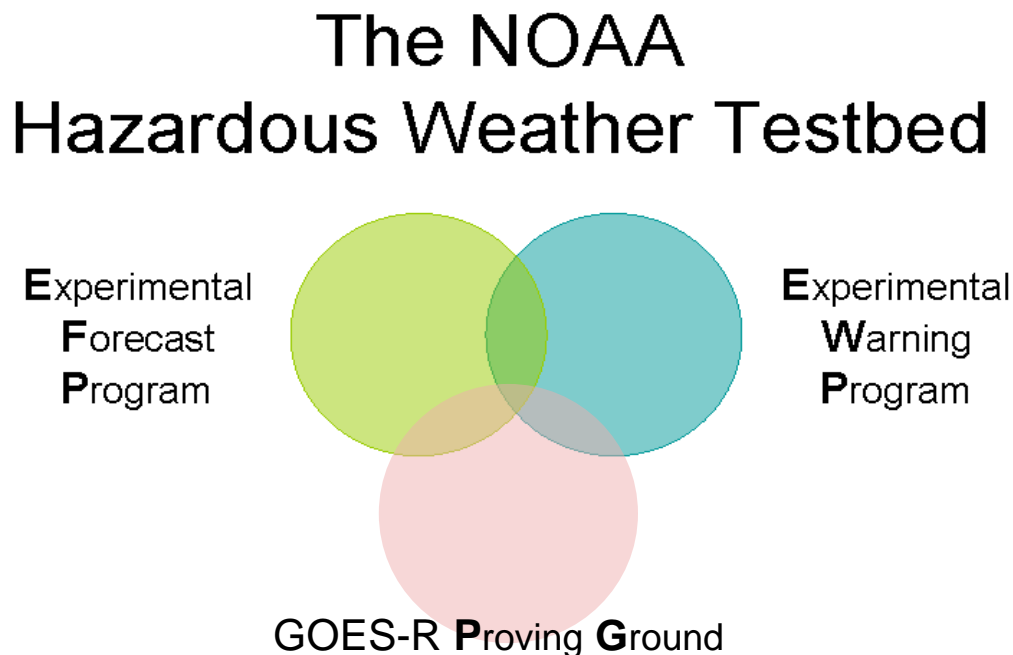


Figure 1: The umbrella of the NOAA Hazardous Weather Testbed (HWT) encompasses two program areas: The Experimental Forecast Program (EFP), the Experimental Warning Program (EWP), and the GOES-R Proving Ground (GOES-R).

As the distinction between warnings and short-term forecasts of convective weather gradually diminishes, the degree of overlap is expected to increase. Both programs reside beneath the overarching HWT organization with a focus on national hazardous weather needs.

In 2009 a GOES-R Proving ground was established at the SPC to test prototype satellite products from the next generation of geostationary satellites. The mission of the Proving Ground encompasses both warning and forecasting applications for hazardous mesoscale weather and testing and validation activities occur in the EFP and EWP parts of the HWT.

The specific mission of each HWT program branch is:

The Experimental Forecast Program - EFP

The EFP branch of the HWT is focused on predicting hazardous mesoscale weather events on time scales ranging from a few hours to a week in advance, and on spatial domains ranging from several counties to the CONUS. The EFP embodies the collaborative experiments and activities previously undertaken by the annual SPC/NSSL Spring Experiments. For more information about the EFP see <http://www.nssl.noaa.gov/projects/hwt/efp/>.

The Experimental Warning Program – EWP

The EWP branch of the HWT is concerned with detecting and predicting mesoscale and smaller weather hazards on time scales of minutes to a few hours, and on spatial domains from several counties to fractions of counties. The EWP embodies the collaborative warning-scale experiments and technology activities previously undertaken by the OUN and NSSL. For more information about the EWP see <http://www.nssl.noaa.gov/projects/hwt/ewp/>.

The GOES-R Proving Ground – GOES-R PG

The GOES-R PG exists to provide pre-operational demonstration of new and innovative products as well as the capabilities available on the next generation GOES-R satellite. The overall goal of the Proving Ground is to provide day-1 readiness once GOES-R launches in late 2015. The PG interacts closely with both product developers and NWS forecasters. More information about GOES-R PG is found at http://cimss.ssec.wisc.edu/goes_r/proving-ground.html.

Rapid science and technology infusion for the advancement of operational forecasting requires direct, focused interactions between research scientists, numerical model developers, information technology specialists, and operational forecasters. The HWT provides a unique setting to facilitate such interactions and allows participants to better understand the scientific, technical, and operational challenges associated with the

prediction and detection of hazardous weather events. The HWT allows participating organizations to:

- Refine and optimize emerging operational forecast and warning tools for rapid integration into operations
- Educate forecasters on the scientifically correct use of newly emerging tools and to familiarize them with the latest research related to forecasting and warning operations
- Educate research scientists on the operational needs and constraints that must be met by any new tools (e.g., robustness, timeliness, accuracy, and universality)
- Motivate other collaborative and individual research projects that are directly relevant to forecast and warning improvement

For more information about the HWT, see <http://www.nssl.noaa.gov/hwt/>

2. Historical Perspective

Co-location of the Storm Prediction Center (SPC) with the National Severe Storms Laboratory (NSSL), the Oklahoma City/Norman Weather Forecast Office, and many University of Oklahoma meteorological organizations in the National Weather Center in Norman provides a unique opportunity to enhance long-standing community interactions and collaboration on a variety of operationally relevant research and experimental forecast programs. Since the re-location of the SPC to the previous NSSL facility Norman in early 1997, a wide cross section of local and visiting forecasters, research scientists, and model developers has participated in a number of experimental programs since the late 1990s. These include forecasting support for field programs such as the International H2O Project (IHOP) and VORTEX2, establishing the SPC winter weather mesoscale discussion product, evaluating operational and experimental NWP models for application in convective forecasting, including Short Range Ensemble Forecast (SREF) systems and convection-allowing Weather Research and Forecasting (WRF) models, and integrating new observational data, objective analyses, and display tools into forecast operations.

A key goal of these programs is to improve forecasts of hazardous meteorological phenomena by: 1) accelerating the transfer of new technology and research ideas into forecast operations at the SPC and other NWS offices, and 2) sharing new techniques, skills, and applied research results more freely with others in the operational forecasting community. Typical issues addressed in these activities include, but are not limited to: optimizing use of vast and ever increasing quantities of observational and model data in operational forecasting, testing and evaluation of new NWP models, better understanding of operational forecast problems, development and evaluation of diagnostic conceptual models, and new product development and display strategies utilizing operational workstations.

Each spring during the climatologically most intense severe weather period, annual multi-agency collaborative forecasting experiments known as the HWT EFP Spring Experiment (formerly called the SPC/NSSL Spring Program) have occurred since 2000. The only exception was in 2006 when the physical move to the new National Weather Center building

precluded a large collaborative experiment. During that spring SPC conducted a focused internal pre-implementation evaluation of the NCEP NAM-WRF model.

The following sections provide additional background information about the motivation for the Spring Experiments, the SPC national severe weather forecasting mission and associated scientific and service challenges, an overview of the scientific goals of the 2010 Spring Experiment and its relevance to operational forecasting, the schedule of daily forecasting and evaluation activities, and a list of weekly participants for the 2010 Spring Experiment.

3. Spring Experiment Background and Motivation

Operational Forecasting of Severe Convective Storms: Current State and Challenges

The prediction of convective weather is important from both meteorological and public service/societal impact perspectives. A primary mission of the National Weather Service is the protection of life and property from hazardous weather phenomena, and applied research aimed at improving the prediction of high impact weather such as severe thunderstorms and tornadoes is a critical activity at the NSSL, SPC, OUN, and other NWS offices.

The SPC is responsible for the prediction of severe convective weather over the contiguous United States on time scales ranging from several hours to eight days. To meet these responsibilities, the SPC issues Convective Outlooks for the Day 1, Day 2, Day 3, and Day 4-8 periods to highlight regions with enhanced potential for severe local storms (defined as thunderstorms producing hail ≥ 1 inch in diameter, wind gusts ≥ 50 kt or thunderstorm induced wind damage, and/or tornadoes). These Outlooks are issued in both categorical (slight, moderate, or high risk) and probabilistic formats, using graphical and text products, and are issued with increasing frequency as the severe weather time frame draws nearer. In addition to the scheduled Outlooks, Severe Thunderstorm and Tornado Watches are issued as needed to provide a higher level of alert over smaller regions in time and space when atmospheric conditions are favorable for severe thunderstorms and/or tornadoes to develop. The SPC also issues short-term Mesoscale Discussion products that emphasize hazardous weather on the mesoscale and often serve to fill the gap between the larger scale Outlooks and smaller scale Watches.

The suite of specialized hazardous weather forecast products depends on the ability of SPC forecasters to assess the current state and evolution of the environment over varied time frames, and to synthesize a wide variety of observational and numerical model data sources. In general, observational data play a dominant role in diagnostic assessment for short-term forecasting, however, the development of more accurate and higher resolution models in recent years has allowed model information to influence the short-term prediction of convection as well. This is especially evident in the use of the hourly Rapid Update Cycle model, which forms a foundation for the SPC Mesoscale Analysis fields.

An effective NWS severe weather forecast and warning program should provide the public and other specialized users with sufficient advance notice of impending hazardous weather. Human response studies have shown that when a severe thunderstorm or tornado warning is issued, people are more likely to seek safe shelter if they have been made aware of the severe weather threat prior to the issuance of the warning. However, if they have not been pre-

conditioned to the threat prior to hearing a warning, their first response is often to seek confirmation of the threat, rather than to seek shelter. This can result in the loss of critical reaction time when life and property are at immediate risk. Thus, there is a substantial need for the SPC to issue severe weather watches prior to the issuance of warnings by local NWS Weather Forecast Offices (WFOs), in order to allow WFO staffs, emergency managers, broadcast media, etc. sufficient time to implement contingency plans prior to the onset of severe weather.

The Need for More Detailed Thunderstorm Forecasts

This goal places additional requirements on SPC forecasters to determine in advance the characteristics of potential severe thunderstorm activity. Operational experience and research studies suggest that the type of severe weather that occurs (tornadoes, hail, or damaging winds) is often closely related to the convective mode (or morphology) exhibited by storms, such as discrete cells, squall lines (or quasi-linear convective systems -QLCS), and multi-cellular convective systems. A disproportionate number of intense tornado and widespread straight-line wind damage events appear to be associated with two dynamically unique classes of thunderstorms: supercells and bow echoes. Thus, accurate severe weather watches are dependent on forecasters being able to predict properly not only where and when severe thunderstorms will develop and how they will evolve over the next 2 – 8 hours, but also the convective mode(s) that are most likely to occur.

There is also an increasing requirement to provide higher temporal resolution forecast information on thunderstorms and a variety of associated hazardous weather phenomena, including severe local storms, heavy rain/flash flooding, lightning strike potential, and aviation-related hazards of turbulence, icing, and low-level wind shear. Users such as emergency managers and other first responders, air traffic flow managers and others in transportation, power companies, etc., need greater time/space specificity in thunderstorm forecasts. The SPC is now providing higher temporal resolution thunderstorm forecasts on an experimental basis (see <http://www.spc.noaa.gov/products/exper/enhtstm/>), in part, to support aviation forecasters at the NOAA/NWS/NCEP Aviation Weather Center (AWC) and air traffic managers at the FAA Air Traffic Control System Command Center (ATCSCC).

Given the SPC's primary mission of mesoscale forecast responsibility, we continue to place a strong emphasis on assessing the current state of the atmosphere by using real-time observational data and derived diagnostic parameters for short-term thunderstorm prediction. The GOES-R PG plays a key role in developing and demonstrating real-time satellite-based diagnostic and short-term prediction products for use in convective forecasting. However, owing to insufficient sampling of the mesoscale environment (especially when the horizontal and vertical distribution of water vapor is considered) coupled with limited scientific knowledge of important mesoscale and storm-scale processes, considerable uncertainty exists in the prediction of deep convection. While traditional operational models such as the NAM and GFS often can predict broader regions of precipitation utilizing parameterized convection, they are not capable of resolving important details of the smaller scale convective structure that are critical to severe weather forecasters. Furthermore, various proximity sounding studies using observed radiosondes and RUC model analyses indicate that the relationship between environmental characteristics (such as CAPE and vertical shear) and storm mode is not unique; rather it is found that similar storm types occur within different parts of the CAPE-shear parameter space, and different storm types occur within similar parts

of parameter space. Therefore, in recent years the Spring Experiment has been focusing on testing and evaluating cutting edge high resolution convection-allowing NWP models to determine potential contributions to operational severe weather forecasting.

Evaluation of Deterministic Convection-Allowing NWP in the Spring Experiment

Earlier research studies using idealized cloud resolving models to simulate deep convective storms at the National Center for Atmospheric Research (NCAR) and the University of Oklahoma Center for Analysis and Prediction of Storms (CAPS), among others, indicated that in some cases the models could replicate severe storm structures including supercells and bow echo systems. However, it was not until recently that sufficient computer resources, communications bandwidth, and advanced workstations became available to facilitate the testing of convection-allowing WRF model configurations over large domains in a semi-operational forecasting environment, and to assess their potential utility for severe weather forecasting. It has been demonstrated over the last seven years through Spring Experiments, field programs such as BAMEX and VORTEX2, and daily use by SPC forecasters of 4 km WRF models from the NCEP Environmental Modeling Center (EMC) and NSSL, that convection-allowing configurations of the WRF model can predict convective storms that, at times, appear remarkably similar to actual storms as seen on radar.

Progress has also been made in developing output fields such as simulated reflectivity that displays model-generated precipitation systems and storms that are visually similar to radar-derived images of actual storms. This allows forecasters to apply their knowledge of storm structure, intensity, and associated severe weather threats gained through observation of radar detected storms to aid in their interpretation of model generated storms. Furthermore, extraction of new parameters such as updraft helicity (a marker for a rotating updraft) has benefited forecasters by identifying explicit storm attributes that indicate enhanced severe potential. This is in contrast to traditional approaches where forecasters utilize mesoscale model output to provide information about evolution of the pre-convective environment, and then they use their knowledge of model biases and thunderstorm physical processes to determine the spectrum of storms that are possible. The first generation of operationally applied convection-allowing models takes this one step further, as they provide explicit information about the types of storms that may develop within predicted mesoscale environments.

Experiments with different WRF model configurations also indicate that it is not uncommon for each of the models to produce a variety of convective solutions for initiation, mode, and evolution, especially within more weakly forced environments. Thus, the model forecasts appear to reflect various uncertainties associated with real-world convective forecasting. These uncertainties arise primarily from: 1) the need to better sample and predict the pre-convective and near-storm environments, as deep convection can be sensitive to small variations in the mesoscale environment, and 2) limits in our understanding of smaller scale physical processes relevant to convection, which are modulated by mesoscale and storm-scale forcing that are difficult to assess in the actual atmosphere.

Several years of experience with 00 UTC “cold start” WRF models using NAM model initial conditions and lateral boundary conditions (ICs/LBCs) have also revealed that it takes several hours of “spin up” time before the models can generate coherent, stable precipitation systems. These “cold start” runs are typically unable to provide substantial short-term

guidance in the 0-6 hr time frame, but they have often demonstrated value in providing useful guidance for next day's diurnal heating cycle during the 18-30 hr forecast period. It has also been seen that the larger scale forcing provided by the "parent" NAM ICs/LBCs modulates the areas of convective storm development in the WRF models. This is particularly evident within strongly forced environments where the WRF convective storms have a tendency to occur in regions where the NAM generates larger scale areas of precipitation.

If WRF models initialized at 00 UTC are to provide useful forecast guidance for the next day's diurnal heating cycle, they must correctly spin up deep convection during the evening, then predict properly the evolution of the storms and their impact on the environment during the overnight hours. If this sequence of events is poorly represented, the pre-convective environment in the model during the subsequent afternoon may not replicate the actual environment, and the model prediction of storms may reflect errors in the environment specification. For example, if the 00 UTC model forecast erroneously maintains convective storm systems too late into the morning, the effects of precipitation, clouds, and an expanding low-level cold pool/convective outflow may maintain a stable environment that is unfavorable for later storm development. When this type of error occurred during the 2008 Experiment, the model(s) typically underpredicted afternoon storm development in areas where the spurious cold pool was located. On the other hand, when the 00 UTC models predicted correctly the evolution of nocturnal storms, they were much more likely to produce skillful forecasts of storms for the next afternoon and evening.

These findings stress the critical importance of predicting correctly the evolution of the mesoscale environment, and suggest that the ability to run "update" models at later times with new ICs/LBCs can be of value to forecasters. In 2008, the EMC High Resolution Window WRF-NMM runs at 12 UTC were often compared with 00 UTC WRF runs on days when the earlier runs were determined to have predicted inaccurate environmental conditions by late morning (e.g., misplaced surface boundaries and errors in thermodynamic fields). In many of these cases, the 12 UTC update run predicted the afternoon environment more accurately and this translated into improved convective forecasts.

Hourly Maximum Fields from WRF Models

Traditionally, output from numerical weather prediction (NWP) models has been presented to forecasters as a series of snapshots in time (an exception is accumulated precipitation). As model resolution and transmission bandwidth have increased, the time interval between these snapshots has decreased. For most forecasting applications, hourly output is adequate because the evolution of common larger-scale features of interest (i.e., fronts, jet streaks, low and high pressure centers, etc.) is well sampled by the hourly frequency. Furthermore, this frequency is a pragmatic choice because the sheer volume of data associated with more frequent output files would exceed the capacity of current dissemination, processing, and storage systems. However, as NWP applications move to higher resolution, the features of interest begin to change and hourly sampling can become inadequate.

Simulated convective storm features such as reflectivity and updraft helicity (UH) features often evolve on convective time scales commonly measured in minutes, not hours. Thus, it is important to monitor model storm behavior at a higher frequency than

hourly output provides. This rationale is similar to operational monitoring of actual storms using radar, where no one would consider hourly snapshot images of storms to be adequate. Rather than simply outputting model fields on a much more frequent basis, a strategy has been developed to monitor and track small-scale, rapidly changing convective storm features every model time step between regular hourly model output times. The individual grid point temporal maxima during each hour are saved and output at the regular hourly intervals, providing a useful perspective on the maximum intensity and track of strong convective phenomena in the model forecasts.

This data processing is intended to fill in the temporal gaps between the standard top of the hour model output and provide unique information about the most intense storm attributes, which are unlikely to occur only at the hourly output times. Currently, the tracking of “history variables” is applied to low level simulated reflectivity, updraft speed, downdraft speed, updraft helicity, 10 m wind speed, and vertically integrated graupel grids.

The computation of hourly maximum fields was first introduced in the NSSL WRF model, and has been subsequently incorporated into WRF models run by a number of major modeling centers (including EMC, GSD, CAPS, NCAR, and AFWA). These fields will be available from all WRF model configurations used in the Spring Experiment. This approach represents an important first step in exploring ways to extract new output fields and/or compute new diagnostics from convection-allowing models, and the output has been utilized in SPC operations for two years with promising results.

Radar Assimilation into Convection-Allowing Models

To fully capitalize on high resolution models to provide short-term forecast guidance on convective scales, advanced data assimilation techniques that include 3D radar reflectivity and velocity fields are necessary in order for the models to “know” where storms are located at the start of the model run. This very challenging task was introduced into the Spring Experiment in 2008, as CAPS used a real-time 3DVAR system to assimilate radar data over a three-fourths CONUS domain for the first time. Although the impact of the radar assimilation on the model forecasts typically appeared to diminish after several hours, this experimental arena will be a focus of activity in coming years.

NOAA/ESRL/GSD has been developing an experimental 3 km WRF version called the High Resolution Rapid Refresh (HRRR) model for several years, and output from this system will be examined more closely this year. The HRRR is nested within the hourly 13 km backup RUC model, and uses the RUC 3DVAR data assimilation system (including radar data) and the diabatic digital filter initialization (DDFI) procedure. This cycled system creates realistic vertical thermal and convergence-divergence couplets in the model atmosphere based on the presence of radar and lightning indicated convection, which allows improved dynamical balance to support existing convection in the short-term model forecasts. The HRRR uses a 1-hr RUC forecast of reflectivity at the initial time, and taking advantage of the DDFI it is able to downscale from the 13 km RUC to the 3 km HRRR grid within the first 15-30 minutes of the integration.

Convective Predictability on the Grid Scale and Storm Scale Ensemble Forecast System

Our experience has also shown that variations in WRF model convective storm predictions are at times difficult for operational forecasters to reconcile, in part because all solutions may appear to be plausible for a given mesoscale environment. Thus, the forecaster must determine how much confidence to place in specific model solutions, which is often difficult to assess because very high resolution models will attempt to predict phenomena (such as thunderstorms) on scales that are inherently unpredictable. The uncertainty in thunderstorm prediction suggests at least several possible research approaches to explore: 1) development of appropriate data assimilation systems for convection-allowing models to better resolve the initial conditions, and 2) improvement in the model itself with more realistic physics and increased resolution. However, inherent predictability limits at convective scales necessitate application of ensemble forecasting strategies, similar to those currently used operationally for synoptic scale and mesoscale forecasting, to address challenges of convective-scale forecasting.

A Storm-Scale Ensemble Forecast (SSEF) system has been tested in Spring Experiments since 2007 to systematically explore aspects of uncertainty in thunderstorm prediction. Although questions remain concerning appropriate perturbation strategies for a convection-allowing ensemble system, experiments with 10-20 member SSEF systems in 2007-2009 have shown promising results. As computing resources have expanded, the SSEF evolved from a 10 member WRF-ARW ensemble in 2007 that contained 5 mixed-physics only members and 5 mixed-physics + perturbed IC/LBC members, to a 20 member multi-model (ARW, NMM, ARPS) ensemble in which 17 members contained both mixed physics and IC/LBC perturbations in 2009. Further refinement of the SSEF will occur this year as additional members containing more sophisticated physics perturbations are added to the system. In addition, development of new display tools for probabilistic assessment of thunderstorm potential and model-generated storm characteristics utilizing “neighborhood” approaches that more properly reflect limits to grid scale predictability will enhance our ability to utilize SSEF output.

This work links directly toward the Warn-on-Forecast (WoF) concept that envisions the use of an ensemble system that assimilates observations of convective storms and their environments into high-resolution, explicit convective-scale numerical weather prediction models, creating a probabilistic convective scale analysis and forecast system. The SSEF system includes all the fundamental components of WoF - e.g., assimilation of radar and environmental data, model configurations capable of explicitly resolving deep convective storms (no parameterized convection), and ensemble-based probabilistic prediction - but all of these components will need considerable improvement before WoF reaches fruition. However, our experience over the past few years suggests that progress in convective scale probabilistic prediction is occurring first on somewhat larger time/space scales compared to true WoF, and the SSEF is a logical first step in this direction.

Finally, a key component of the annual experiments is the participation of operational forecasters from the SPC, other NCEP Centers, NWS WFOs, Environment Canada, and several private sector companies. Their insights and experience provide a real-world severe weather forecasting perspective when assessing the usefulness of convection-allowing WRF modeling systems, and provide them with opportunities to become familiar with cutting-edge science and technology applications before they are implemented operationally. This

operational-research link increases the likelihood that HWT activities will result in improved severe weather forecasts and better public service. Forecaster interactions with model developers, research scientists, university faculty, and graduate students create a unique forum where a diverse mix of scientific backgrounds and insights work together to advance operationally relevant research and improve forecasts of hazardous convective weather.

4. Additional HWT Focus Areas in 2010 – Aviation-Impacts and QPF

Convective storms have a wide variety of societal impacts that range beyond the traditional HWT focus on severe convective weather. According to NOAA economic statistics, warm-season thunderstorms cause ~70% of air traffic delays in the U.S. and cost the economy upwards of \$4 billion dollars each year. In addition, thunderstorm-generated heavy rain and flash floods are one of the leading causes of weather-related fatalities averaging ~130 deaths per year. Improved forecasts of thunderstorms will result in large societal benefits, and it is appropriate for the HWT to explore additional thunderstorm hazards during the Spring Experiment. In 2010 we are collaborating with the AWC and the HPC to examine the ability of convection-allowing models to provide useful guidance for thunderstorm-aviation impacts and heavy rain forecasting.

Applications to Thunderstorm Impacts on Warm-Season Aviation Operations

To assist the AWC in the development of their Aviation Weather Testbed, the HWT Spring Experiment will include an aviation-impacts component to complement the traditional HWT focus on severe convection. Thunderstorms are responsible for many air traffic delays across the National Airspace System (NAS) each year. They are also considered a threat to aviation safety due to their ability to produce both en route and terminal weather-related hazards such as lightning, hail, turbulence, microbursts, and low-level wind shear. The Federal Aviation Administration (FAA) plans traffic flow management (TFM) to avoid thunderstorms utilizing a 6 hour forecast product designed for aviation. The AWC plays a key role in providing detailed short-term thunderstorm forecasts through the Collaborative Convective Forecast Product, which is issued every two hours and provides thunderstorm forecasts valid at 2 hr intervals out to 6 hrs that include information on storm coverage, growth rates, movement, thunderstorm tops, presence of thunderstorm lines, and forecaster confidence (for more information about the CCFP see <http://aviationweather.gov/products/ccfp/>).

Convection-allowing WRF model output has potential to provide very detailed hourly forecasts of convective storms, especially related to storm coverage and mode, and can provide forecasters and aviation traffic flow managers with potentially useful information about future storm impacts on en route aircraft as well as threatening storm conditions near hub airports. In more strongly forced situations lines may be more confidently predicted, but small errors in timing on the order of 1-2 hrs can create large disruptions in the NAS. The generally limited predictability of storms on the grid scale must be acknowledged when utilizing high resolution WRF model forecasts of storms, and results from previous years suggests that while deterministic forecasts may be compatible with historical TFM practices, the forecast process is better suited to probabilistic convective

weather information given the uncertainty in predicting exact times and locations of thunderstorms.

Thus, the use of SSEF output fields in combination with traditional mesoscale model output (NAM, SREF), deterministic convection-allowing WRF model forecasts, and statistical thunderstorm guidance from the Local Aviation MOS Program (LAMP) will be tested and evaluated during the 2010 Spring Experiment. There are also plans to utilize experimental output from the Consolidated Storm Prediction for Aviation (CoSPA) project if that data become available for use during the latter part of the Spring Experiment. Aviation forecasters from AWC and NWS offices will work with traffic managers from the FAA ATCSCC in the HWT aviation-impacts component to assess the ability of the new guidance to improve thunderstorm forecasts for aviation interests. The primary focus will be on afternoon thunderstorms over the eastern half of the CONUS, especially over the northeast corridor where storm disruptions can impact traffic flow across large parts of the country. Finally, to begin addressing a new strategic planning initiative to provide next day guidance on potential thunderstorm impacts, an experimental Day 2 aviation thunderstorm forecast will be issued each afternoon as part of a Collaborative Strategic Planning Process at the FAA ATCSCC.

Applications to Quantitative Precipitation Forecasting

To assist the HPC in the development of their Hydrometeorological Testbed (HMT), the HWT Spring Experiment will also include a QPF component to complement the traditional HWT focus on severe convection. It has been long noted that QPF scores exhibit lower skill during the warm season, and this is largely attributable to the dominant contribution from convection on warm season precipitation. Traditional synoptic scale and mesoscale NWP models such as the GFS and NAM use convective parameterization schemes (CPS) to account for the sub-grid scale effects of deep convection, and the CPS have tendencies to exhibit a number of systematic errors. These include: erroneous precipitation “bulls-eyes”, considerable phase errors in time and space, especially for MCS development that accounts for much of the warm season rainfall across the US, and a low bias for the most critical heavy rain producing thunderstorm events. Previous studies have found that convection-allowing models have the ability to better predict convective mode, provide more realistic amplitude of rainfall, and better represent the diurnal cycle and propagation of rainfall systems. It has also been demonstrated that a SSEF with a relatively small number of members has improved QPF skill compared to a larger mesoscale ensemble using parameterized convection.

For an initial test and evaluation, the QPF forecast teams will incorporate guidance from convection-allowing WRF models including the SSEF to produce experimental probabilistic QPF forecasts for 6 hr periods valid 18-00 UTC and 00-06 UTC that cover the primary diurnal convective storm periods. (It is recognized that the climatological nocturnal precipitation maximum over the plains during the warm season may occur after 06 UTC, but the SSEF forecast period ends at 06 UTC and restricts the experimental forecasts to the 18-06 UTC period.) Forecasters will use the experimental model guidance to supplement traditional model guidance (e.g., NAM, GFS, SREF) in the

forecasting process. The experimental forecasts will depict contours for the probability of exceeding (POE) 0.5" and 1" thresholds for each 6 hr period, using categorical terms of slight = 25%, moderate = 50% and high = 75% probability values. In addition, to explore the utility of the convection-allowing models to better predict localized heavier precipitation amounts, each forecast that includes a probability of 1" or greater will also identify the maximum value within the 1" POE for each 6 hr period.

The HWT activities will occur in conjunction with an initial in-house experiment at the HPC to familiarize QPF forecasters with the experimental model output, and to begin assessing the challenges as well as the potential value and utility of convection-allowing model guidance for QPF application.

5. Developmental Testbed Center Objective Evaluation Background

New Objective Verification Approaches

Subjective verification of model forecasts has been a cornerstone to HWT activities in previous years. This approach has provided valuable insights into how forecasters use numerical models, and facilitates the gathering of information about the value of new guidance tools from the perspective of a forecaster. In addition, traditional verification measures (e.g., Equitable Threat Score or ETS) used for synoptic scale and mesoscale model forecasts of discontinuous variables such as precipitation typically provide less useful information (and even misleading information) about forecast accuracy as the scale of the phenomena being evaluated decreases. This is because the ETS is proportional to the degree of grid scale overlap in space and time between the forecasts and observations, and there is typically low predictability on convective scales. Despite these limits, operational severe weather forecasters have often found value in WRF forecasts of thunderstorms and convective systems, since they can provide unique information about convective mode, coverage, and evolution that is not resolved by mesoscale models using parameterized convection. In recent years, we have found that subjective evaluation has great potential to serve as a comparative benchmark for assessing new objective verification techniques designed for high resolution NWP, and has had a significant positive impact on model development strategies.

In order to better utilize subjective and objective verification techniques in a complementary manner, simulated composite reflectivity and 1-hr QPF output from several model runs will be evaluated using subjective visual comparisons and objective statistical measures produced by the Developmental Testbed Center's (DTC) Meteorological Evaluation Tool (MET). The focus this year will be on probabilistic predictions, particularly of extreme precipitation events and severe weather as it relates to aviation weather. All members of the Center for Analysis and Prediction of Storms (CAPS) Storm Scale Ensemble Forecast (SSEF) system will be evaluated for select variables. Ensemble products from the fifteen members selected by the NOAA Storm Prediction Center (SPC) will also be evaluated. Operational (or near-operational) models will be used as a baseline for comparison. These include the North American Mesoscale (NAM) model, the

High Resolution Rapid Refresh (HRRR), and ensemble products from the Short Range Ensemble Forecast (SREF) system. Other contributing models will be brought in and archived for retrospective studies.

MET is designed to be a highly-configurable, state-of-the-art suite of verification tools. We will focus on the use of the object-based verification called Method for Object-based Diagnostic Evaluation (MODE) that compares gridded model data to gridded observations for the QPF and simulated reflectivity forecasts. MODE output will be tested to evaluate its ability to diagnose different types of convective modes considered important in forecasts and observations of convective weather, such as linear systems, discrete cells, and MCS's. Traditional verification statistics will also be computed. Details about the DTC MET system is at <http://www.dtcenter.org/met/users/>.

Verification “truth” will be provided by NSSL National Mosaic and Multi-Sensor QPE (NMQ) multi-sensor Quantitative Precipitation Estimates (QPE) and three-dimensional radar reflectivity data bases. See <http://www.nssl.noaa.gov/projects/q2/> for more information about the NMQ. A list of model fields and verifying datasets that will be utilized during the DTC objective evaluation is found in Table 5, and a thorough summary of DTC verification approaches is found in Appendix G.

Table 5. List of variables (and thresholds) to be evaluated during SE 2010. Many will be available in real-time and others will be generated retrospectively. Traditional and Spatial metrics for which models are also listed.

Models and Fields to be Evaluated

FCST Field	Observation	Grid-Stat	MODE	Models
Prob of Exceed (0.25", 0.5", 1", 2" over 3 and 6 hrs)	0.25", 0.5", 1", 2" QPE over 3 and 6 hrs	Brier Score, Decomp of Briar score, Area under ROC, Reliability Diagram	none	Ensemble products from CAPS and SREF
50% Prob of Exceed (0.25", 0.5", 1", 2" over 3 and 6 hrs)	0.25", 0.5", 1", 2" QPE over 3 and 6 hrs	None	MMI, Intersection Area, Area Ratio, Centroid Distance, Angle Difference, % Objects and Area Matched, 50 th and 90 th Percentile of Variable	Ensemble products from CAPS and SREF
0.25", 0.5", 1.0", 2" QPF over 3 and 6 hrs	0.25", 0.5", 1.0", 2" QPE over 3 and 6 hrs	GSS, CSI, FAR, PODY, FBIAS	Same as above	CAPS members, CAPS ens mean, SREF ens mean, HRRR, NAM
Sim. CompositeRefl (20,30,40,50 dBZ)	Q2 Composite refl (20,30,40,50 dBZ)	GSS, CSI, FAR, PODY, FBIAS	Same as above	CAPS members, CAPS ens mean, SREF ens mean, HRRR, NAM
18 dBZ Echo Top (18, 25, 30, 35, 40, 45 kft)	Q2 18dBZ Echo Top (18, 25, 30, 35, 40, 45 kft)	GSS, CSI, FAR, PODY, FBIAS	Same as above	CAPS members, CAPS ens mean, SREF ens mean, HRRR, NAM
Prob of 40dBZ echos	Q2 Composite reflectivity (40dBZ)	GSS, CSI, FAR, PODY, FBIAS		Ensemble products from CAPS and SREF
50% Prob of 40dBZ echos	Q2 Composite reflectivity (40dBZ)		MMI, Intersection Area, Area Ratio, Centroid Distance, Angle Difference, % Objects and Area Matched, Diff between objs of 50 th and 90 th Percentile	Ensemble products from CAPS and SREF

6. Experimental Models

The 2010 Spring Experiment will benefit from the continued participation and key contributions from CAPS, EMC, GSD, and NCAR. Each of these collaborators (along with NSSL) will generate high resolution, convection-allowing model guidance initialized at 00 UTC, and most will provide additional model runs at 12 UTC and/or other times during the convective day. Model domains will generally cover from three-fourths to full CONUS regions, and most 00 UTC models will produce forecasts to at least 30 hrs. CAPS will also produce a subset of WRF runs at 09, 12, 15, and 18 UTC which will focus on the VORTEX2 field program domain centered over the plains states and generally provide forecasts out to 06 UTC, while the hourly GSD HRRR runs which will go out to 15 hrs.

CAPS Models

4 km Storm Scale Ensemble Forecast, 1 km WRF-ARW, and VORTEX2 4 km WRF-ARW runs at 09, 12, 15 and 18 UTC

A major CAPS contribution will be a 26 member *Storm Scale Ensemble Forecast (SSEF) system* with grid spacing of 4 km and forecasts to 30 hrs, utilizing the resources of the National Institute for Computational Sciences (NICS)/University of Tennessee located at Oak Ridge National Laboratory. The SSEF will be a multi-model ensemble with 19 ARW members, 5 NMM members, and 2 ARPS members. The SSEF will draw additional initial condition (IC) and physics diversity from mixed IC/physics perturbations in 15 members (12 ARW and 3 NMM), with new physics diversity provided in many ARW members through the introduction of new two new PBL and three new double moment microphysics schemes.

In all members, the background initial condition will come from interpolation of the 12 km NAM analysis. Mesoscale atmospheric perturbations will be introduced in the initial and lateral-boundary conditions of 10 ARW and 3 NMM members by extracting perturbations from EMC's operational Short Range Ensemble Forecast (SREF) system and applying them to the 13 members. In addition, random/recursive perturbations are applied to 3 ARW members. Convective-scale perturbations will be introduced in the initial conditions of 23 members by assimilating reflectivity and velocity data from the national NEXRAD radar network and a cloud analysis as part of a CAPS 3DVAR system. For the remaining two ARW, two NMM, and two ARPS members, identical model configurations will be used for each pair and there will be no other IC or physics perturbations applied. Radar data will be assimilated into one of the two ARW, NMM, and ARPS members (the C0 control member), but not the other (Cn member). Comparison of output from these two pairs of ARW, NMM, and ARPS members will allow us to isolate the impact of the radar data from other sensitivities at 4 km grid spacing.

Overall, the SSEF configuration builds upon lessons learned from the earlier SSEF systems tested during the 2007-2009 Spring Experiments, and the development this year of a larger multi-model, multi-physics, multi-IC SSEF over a full CONUS domain is expected to be more robust and contain improved statistical performance. For operational forecasting applications, it is anticipated the SSEF will provide improved probabilistic guidance on high

impact convective weather events by quantifying aspects of uncertainty and offering further insights about a possible range of solutions.

CAPS will also provide a CONUS domain *single WRF-ARW forecast at 00 UTC with a 1-km grid length* integrated to 30 hrs. Radar data will also be assimilated into the 1 km ARW but there will be no SREF-based perturbations. This will allow a direct comparison with the SSEF ARW control member and a clean measure of sensitivity to 1 versus 4 km grid spacing when radar data are assimilated. Statistical verification measures from Spring Experiments in 2007 and 2008 indicated similar forecast results from the 2 and 4 km ARW forecasts, suggesting that the benefit gained by increasing horizontal resolution was not sufficient to justify the approximate eight-fold increase in computational resources to produce the 2 km run. However, results from 2009 comparing the 1 km run with the SSEF indicate that 1 km ETS and Bias scores for precipitation forecasts were generally superior to all members as well as the probability matched ensemble mean. This is consistent with other high resolution modeling studies that found more realistic convective storms in terms of structure, size, and number of storms beginning to appear when the grid spacing approaches 1 km.

Finally, CAPS will produce two *4 km WRF-ARW runs initialized at 09, 12, 15, and 18 UTC* over a smaller domain centered on the VORTEX2 field program in the plains states. These series of two runs will be integrated to 18 hrs, and they will have physics configurations identical to the two SSEF ARW members without IC/LBC perturbations (one with radar assimilation and one without radar). These runs initialized every three hours are designed to provide updated convective-scale guidance for afternoon and evening storms based on later initial conditions, and utilize resources at the University of Oklahoma Supercomputing Center for Education and Research.

The CAPS computational domain for the 00 UTC SSEF and 1 km WRF is in Fig. 2, and the WRF-ARW “VORTEX2” domain is in Fig. 3. The SSEF member configuration is provided in Tables 1, 2, and 3.

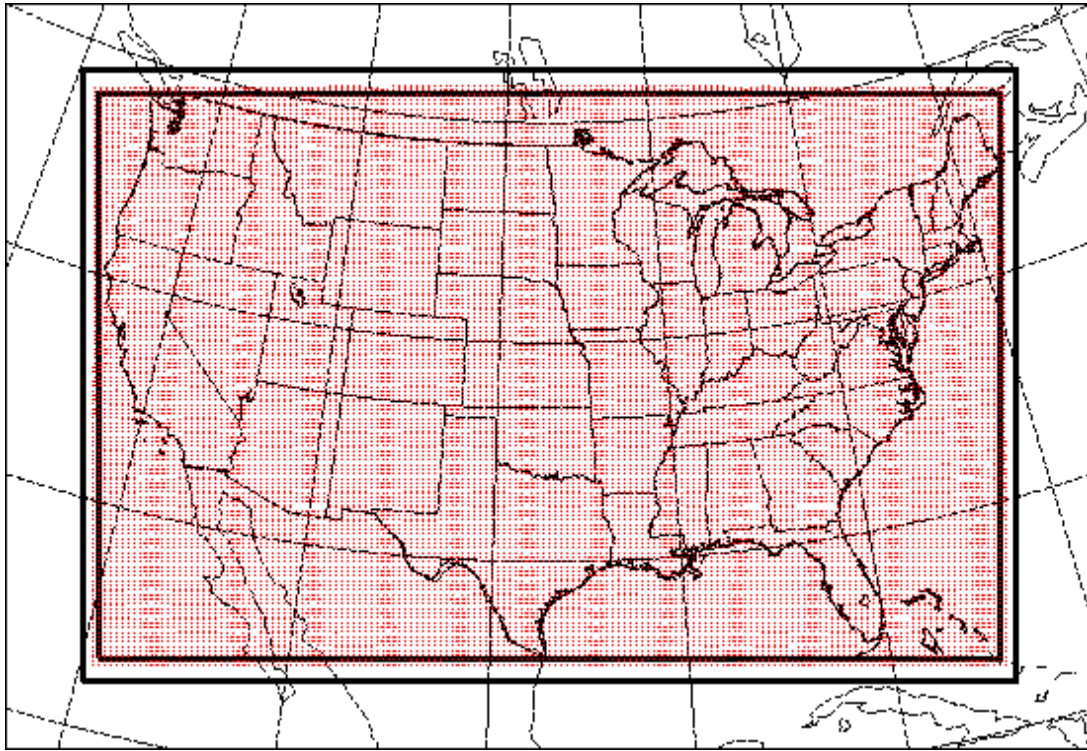


Figure 2. Computational domains for the 2010 Season. The outer thick rectangular box represents the domain for performing 3DVAR (**Grid 1** – 1200×780). The red dot area represents the WRF-NMM domain (**Grid 2** – 790×999). The inner thick box is the domain for WRF-ARW and ARPS and also for common verification (**Grid3** - 1160×720 at 4 km grid spacing; 4640×2880 at 1 km grid spacing).



Figure 3. Model domain used for CAPS 09, 12, 15, and 18 UTC 4 km WRF-ARW update forecasts with 444×480 horizontal grid points.

Table 1. Configurations for ARW members. NAMa and NAMf refer to 12 km NAM analysis and forecast, respectively. ARPSa refers to ARPS 3DVAR and cloud analysis

member	IC	BC	Radar data	microphy	LSM	PBL
arw_cn	00Z ARPSa	00Z NAMf	yes	Thompson	Noah	MYJ
arw_c0	00Z NAMa	00Z NAMf	no	Thompson	Noah	MYJ
arw_m3	arw_cn + random pert	00Z NAMf	yes	Thompson	Noah	MYJ
arw_m4	arw_cn + recursive pert	00Z NAMf	yes	Thompson	Noah	MYJ
arw_m5	arw_cn + em-p1 + recur pert	21Z SREF em-p1	yes	Morrison	RUC	YSU
arw_m6	arw_cn + em-p1_pert	21Z SREF em-p1	yes	Morrison	RUC	YSU
arw_m7	arw_cn + em-p2_pert	21Z SREF em-p2	yes	Thompson	Noah	QNSE
arw_m8	arw_cn – nmm-p1_pert	21Z SREF nmm-p1	yes	WSM6	RUC	QNSE
arw_m9	arw_cn + nmm-p2_pert	21Z SREF nmm-p2	yes	WDM6	Noah	MYNN
arw_m10	arw_cn + rsmSAS-n1_pert	21Z SREF rsmSAS-n1	yes	Ferrier	RUC	YSU
arw_m11	arw_cn – etaKF-n1_pert	21Z SREF etaKF-n1	yes	Ferrier	Noah	YSU
arw_m12	arw_cn + etaKF-p1_pert	21Z SREF etaKF-p1	yes	WDM6	RUC	QNSE
arw_m13	arw_cn – etaBMJ-n1_pert	21Z SREF etaBMJ-n1	yes	WSM6	Noah	MYNN
arw_m14	arw_cn + etaBMJ-p1_pert	21Z SREF etaBMJ-p1	yes	Thompson	RUC	MYNN
arw_m15	arw_cn	00Z NAMf	yes	WDM6	Noah	MYJ
arw_m16	arw_cn	00Z NAMf	yes	WSM	Noah	MYJ
arw_m17	arw_cn	00Z NAMf	yes	Morrison	Noah	MYJ
arw_m18	arw_cn	00Z NAMf	yes	Thompson	Noah	QNSE
arw_m19	arw_cn	00Z NAMf	yes	Thompson	Noah	MYNN

Table 2. Configurations for each individual member with NMM core

member	IC	BC	Radar data	mp_phy	lw_phy	sw-phy	sf_phy
nmm_cn	00Z ARPSa	00Z NAMf	yes	Ferrier	GFDL	GFDL	Noah
nmm_c0	00Z NAMa	00Z NAMf	no	Ferrier	GFDL	GFDL	Noah
nmm_m3	nmm_cn + nmm-n1_pert	21Z SREF nmm-n1	yes	Thompson	RRTM	Dudhia	Noah
nmm_m4	nmm_cn + nmm-n2_pert	21Z SREF nmm-n2	yes	WSM 6-class	RRTM	Dudhia	RUC
nmm_m5	nmm_cn + em-n1_pert	21Z SREF em-n1	yes	Ferrier	GFDL	GFDL	RUC

* For all members: *pbl_physics*=MYJ; *cu_physics*= NONE

Table 3. Configurations for each individual member with ARPS

member	IC	BC	Radar data	Microphy.	radiation	sf_phy
arps_cn	00Z ARPSa	00Z NAMf	yes	Lin	Chou/Suarez	Force-restore
arps_c0	00Z NAMa	00Z NAMf	no	Lin	Chou/Suarez	Force-restore

* For all members: no cumulus parameterization

EMC 4 km WRF-NMM Model

SPC forecasters have used output from earlier versions of the EMC WRF-NMM model since the spring of 2004. The current version is nested within the 12 km NAM and incorporates NAM ICs/LBCs. It is run throughout the year over a CONUS domain Fig. 4) twice daily at 00 and 12 UTC with forecasts to 36 hrs, and output is available to all forecasters via a web page at

<http://www.emc.ncep.noaa.gov/mmb/mpyle/cent4km/conus/00/> . The latter run time provides a morning update for afternoon and evening guidance. In addition, the 12 UTC run will be used to provide initial day 2 guidance for the afternoon aviation experimental forecast for the next day.

Note - if the EMC experimental CONUS WRF-NMM is not available, the operational High Resolution Window 4 km WRF-NMM will be used. This model is also nested

within the 12 km NAM and is initialized twice daily at 00 and 12 UTC over a domain covering the eastern three-fourths of the CONUS, producing forecasts to 48 hrs. The hourly maximum parameters and model forecast soundings are not currently available from the HiResWindow runs.

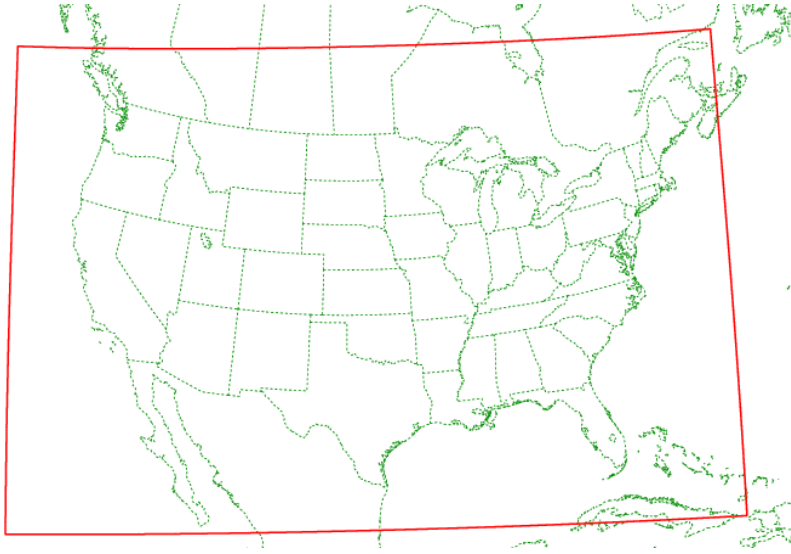


Figure 4. EMC 4 km WRF-NMM domain with 1239x920 horizontal grid points.

NSSL 4 km WRF-ARW Model

SPC forecasters have used output from a 4 km WRF-ARW produced by NSSL since the fall of 2006. This WRF model is run once daily at 00 UTC throughout the year over a full CONUS domain (Fig. 5) with forecasts to 36 hrs. Output is also available on the internet at <http://www.nssl.noaa.gov/wrf/>.

The NSSL-WRF will be outputting several new experimental fields to test and evaluate this spring.

1. Total Lightning Threat (units: flashes km⁻² per 5 min)

There are three total lightning threat experimental parameters that represent microphysical properties of hydrometeor types and charge separation processes within the WRF model convective storms. This is based on work by McCaul and colleagues at SPoRT in Huntsville.

Lightning Threat 1: Upward flux of ice hydrometeors at the -15C level

Lightning Threat 2: Column integrated ice hydrometeors

Lightning Threat 3: Blended solution of Threats 1 and 2 that optimizes temporal variability best depicted by Threat 1 and areal coverage that is best depicted by Threat 2. Threat 3 is very heavily weighted by Threat 1.

These 3 fields are based on the hourly maximum of the ice hydrometeor fields and therefore should be considered to represent the hourly maximum total lightning threats, and are based on work done by McCaul and colleagues at SPoRT. It is recommended that users primarily focus on Lightning Threat 3 field since it statistically combines attributes of the two fundamental physical processes represented in Threats 1 and 2. The Lightning Threat products have only recently been applied to the NSSL-WRF and more work is needed before meaningful quantitative magnitudes can be discerned.

The explicit total lightning is highly dependent on the ability of the NSSL-WRF to predict timing and location of convective storms, but, as with our applications of storm attributes such as UH, this is another step to extract additional storm characteristics from convection-allowing models.

It is anticipated that the output will be most useful for the experimental thunderstorm forecasts issued as part of the aviation-impacts component.

2. Simulated Satellite Imagery

Working with scientists at both CIRA/CSU and CIMSS/UW, simulated satellite imagery is being created from the NSSL-WRF model gridded fields to represent output from a number of channels planned for the GOES-R satellite. The simulated imagery is generated from model gridded surface fields and vertical profiles of predicted moisture, temperature, and clouds, and is sensitive to the microphysics scheme employed in the numerical model.

Selected WRF forecast grids are distributed to both CIRA and CIMSS, where local versions of radiative transfer models are applied to create simulated radiance/brightness temperature fields. The images are then sent to the HWT for display in the N-AWIPS system. CIRA is producing images from 4 infrared channels, and CIMSS is producing output from 8 infrared channels.

This new capability will allow users to directly infer the 4-D evolution of model dynamic processes and associated moisture fields, and to make visual comparisons between satellite observations and operational model output at resolutions comparable to GOES satellite imagery. The simulated GOES imagery allows forecasters to rapidly discern model forecasts of moisture transport, regions of ascent and subsidence, and indications of the vertical extent of clouds including shallow and deep convection. An animated loop of model-derived simulated GOES imagery can allow forecasters and model developers to subjectively ascertain dynamic processes within the model atmosphere very quickly and improve our understanding of model forecast evolution.

The simulated satellite imagery will be available for teams formulating experimental forecasts in the severe storm, aviation-impacts, and QPF components.



Figure 5. NSSL 4 km WRF-ARW domain with 1200x800 horizontal grid points.

GSD 3 km High Resolution Rapid Refresh (HRRR) Model

The 3 km HRRR model is nested within the hourly 13 km RUC model, which provides ICs/LBCs for the HRRR. The HRRR uses a version of the WRF-ARW with generally “RUC-like” physics. A unique aspect of the RUC is the hourly 3DVAR data assimilation system that incorporates a wide array of observational datasets including radar reflectivity via the radar-Diabatic Digital Filter Initialization. The HRRR integration is run over a full CONUS domain (Fig. 6) with forecasts to 15 hrs. At the initial time, the simulated HRRR reflectivity comes from a 1 hr RUC forecast; downscaling from the RUC 13 km grid to the HRRR 3 km grid occurs quickly during the first hour.

The HRRR is being developed to serve users needing frequently updated short-range weather forecasts, including those in the US aviation and severe weather forecasting communities. It is expected to primarily provide guidance for the severe storm and aviation-impacts components of the Spring Experiment, especially during the afternoon updates to the experimental forecasts.

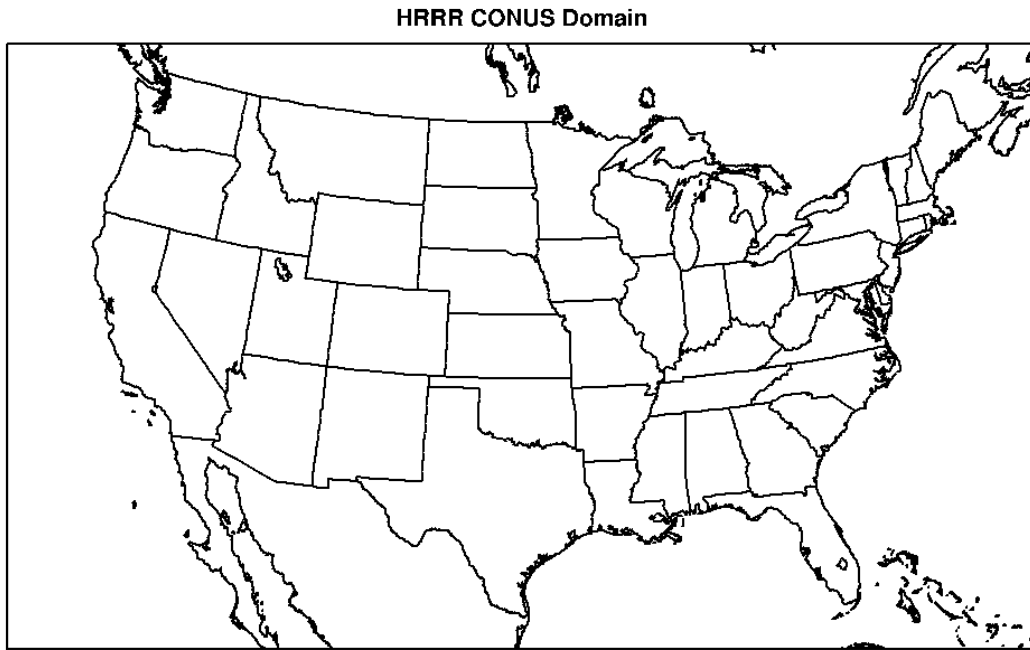


Figure 6. HRRR 3 km domain with 1800x1060 horizontal grid points.

NCAR 3 km WRF-ARW Model

NCAR will focus this spring on running a 3 km WRF-ARW that utilizes initial conditions from the 13 km RUC that includes radar reflectivity via the radar-Digital Filter Initialization. This will use the same ICs that are used by the HRRR, but the LBCs for the NCAR WRF will be provided by the GFS model. The choice of RUC ICs and GFS LBCs will allow further examination of the sensitivity of model forecasts to initial condition specification. It has been noted over the last few years that ICs and forcing attendant to large scale systems provided by the “parent” models played a role in the timing and location of convection in the WRF model forecasts, especially in more strongly forced situations.

The NCAR 3 km WRF-ARW will be run twice daily at 00 and 12 UTC with forecasts out to 48 hrs over a three-fourths CONUS domain (Fig. 7). The 12 UTC run will also be used to provide Day 2 guidance for the afternoon aviation experimental forecast for the next day.

The configuration of the deterministic convection-allowing WRF models for the 2010 Spring Experiment is found in table 4.

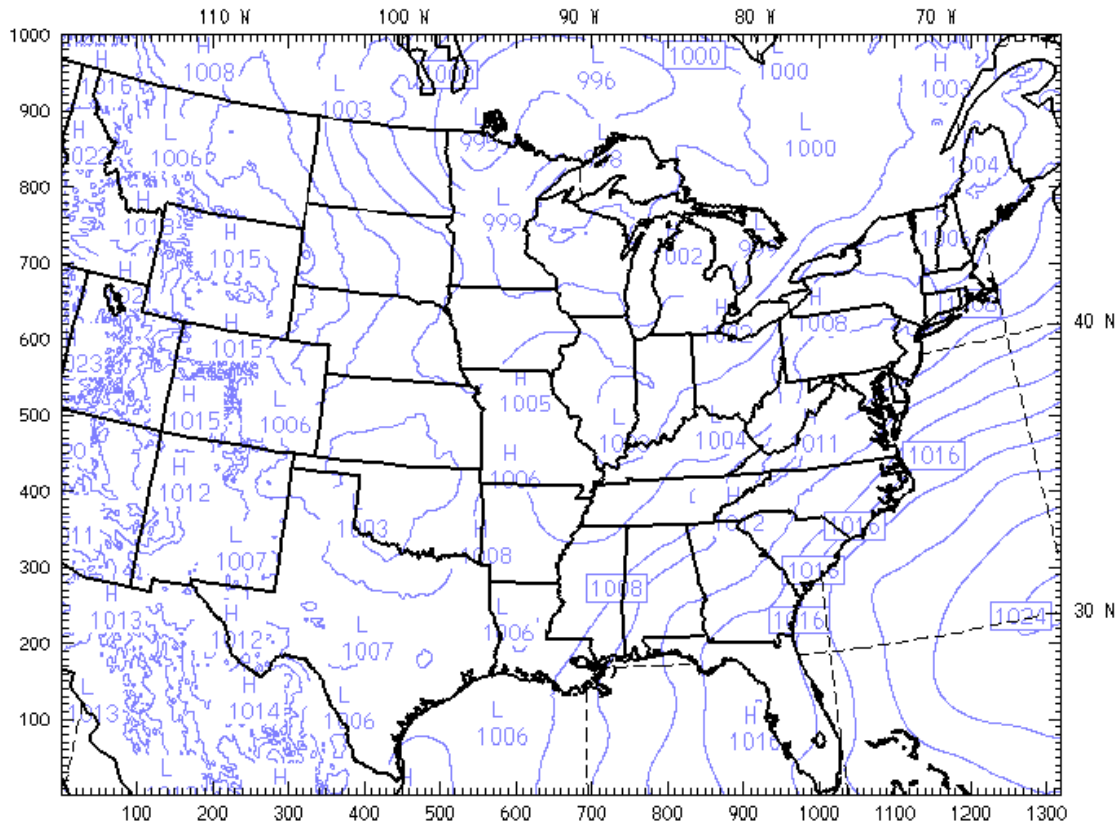


Figure 7. NCAR 3 km WRF-ARW domain with 1320x1000 horizontal grid points.

Table 4. Configurations of deterministic WRF models. The GSD-HRRR3 is initialized hourly with forecasts to 15 hrs; the EMC-NMM4 is initialized at 00 and 12 UTC with forecasts to 36 hrs; the NCAR-ARW3 is initialized at 00 and 12 UTC with forecasts to 48 hrs; the NSSL-ARW4 is initialized at 00 UTC with forecasts to 36 hrs; and the CAPS-ARW1 is initialized at 00 UTC with forecasts to 30 hrs.

	GSD-HRRR3 (ARW)	EMC-NMM4	NCAR-ARW3	NSSL-ARW4	CAPS-ARW1
Horiz. Grid (km)	3.0	4.0	3.0	4.0	1.0
Vertical Levels	50	35	34	35	51
PBL/Turb. Parameterization	MYJ	MYJ	MYJ	MYJ	MYJ
Microphysical Parameterization	Thompson	Ferrier	Thompson	WSM6	Thompson
Radiation (SW/LW)	Dudhia/RRTM	GFDL/GFDL	Goddard/RRTM	Dudhia/RRTM	Goddard/RRTM
Land Surface Model	RUC-Smirnova	Noah	Noah	Noah	Noah
Initial Conditions	13 km RUC	32 km NAM	13 km RUC	40 km NAM	CAPS-3DVAR

7. GOES-R Proving Ground Products

The GOES-R Proving Ground activities will be integrated more directly into the HWT programs this year, as a number of prototype satellite products have applications in the EFP and EWP forecast and warning experiments. The GOES-R visiting scientists will also meet separately for focused discussions in the HWT annex located immediately west of the main HWT facility. The HWT annex is shared with the VORTEX2 Operations Center (VOC). More details about the GOES-R Proving Ground Operations Plan are at: http://hwt.nssl.noaa.gov/Spring_2010/Proving_Ground_Operations_Plan_SPC-FINAL.pdf

The GOES-R Proving Ground products to be demonstrated this year within the Spring Experiment, developing organizations, and HWT focus area include:

- 1) Pseudo-GLM total lightning flash extent and initiation density product (SPoRT/NSSL)
 - Primarily EWP
- 2) CI Nowcast and associated cloud-top cooling rate products (UW-CIMSS/UAH)
 - Primarily EWP but may have near-term EFP nowcast and diagnostic applications
- 3) Overshooting-top and “Enhanced-V” detection product (UW-CIMSS)
 - Primarily EWP but may have near-term EFP nowcast and diagnostic applications
- 4) Simulated GOES-R ABI cloud and moisture imagery using WRF-model microphysics (UW-CIMSS/NSSL and CIRA/CSU/NSSL)
 - Primarily EFP applications for model interpretation
- 5) GOES-R Risk-Reduction products
 - 0-6 hr differential theta-e “Nearcast” product (UW-CIMSS)
 - EFP Short-term forecasting (afternoon updates)
 - 0-3 hr severe hail probability product (CIRA)
 - Possibly EFP very-short term guidance for afternoon updates
- 6) NSSL-WRF Explicit Total Lightning Forecasts (UAH/SPoRT/NSSL)
 - Primarily EFP, especially for aviation and severe applications

8. Objectives and Goals

The primary objectives of Spring Experiment 2010 are listed by component:

A. Severe Convective Storm Component (Leader: SPC)

- Continue test and evaluation of high-resolution convection-allowing models (CAMs) and SSEF to provide useful guidance to severe weather forecasters for high-temporal resolution experimental probabilistic severe weather forecasts. This will focus on improving forecasts of initiation, evolution, mode, and intensity of convective storms.
- Assess the perceived value of probabilistic products and other unique guidance from the SSEF to aid in the forecaster formulation of experimental probabilistic 4-hr severe weather forecasts. These products are designed to complement the current Day 1 experimental enhanced resolution probabilistic thunderstorm products issued by the SPC.
- Explore the utility of GOES-R Proving Ground products as part of an integrated data suite to support severe weather forecaster decision-making.
- Using experimental and operational data sets, determine if forecasters can create reliable probabilistic severe weather products for the occurrence of severe storms (tornadoes, large hail, damaging wind gusts) for two 4-hr periods of 20-00z and 00-04z encompassing the diurnal convective cycle. The forecasts will also include potential for significant severe weather (tornado EF2+, hail $\geq 2''$, wind gust ≥ 65 kt).
- Examine the relationship between model predicted environment information (e.g., CAPE, shear, boundary placement) and model-generated convective storms, and compare skill of environmental predictions from operational mesoscale models (e.g., NAM, SREF) and CAMs/SSEF.
- Examine sensitivity of additional physics diversity in SSEF (e.g., double-moment microphysics and new PBL schemes) on environment/sounding structure and storm development in selected member comparisons.
- Enhance collaboration with the DTC in the development and testing of traditional and new objective measures to provide useful verification of 1-hr simulated composite reflectivity and precipitation forecasts produced by the SSEF radar/no-radar control members and the HRRR model in the 0-12 hr period.
- Build cross cutting relationships between members of the severe weather, aviation, and QPF communities to strengthen collaborations focused on shared thunderstorm forecast challenges.

- Explore forecast consistency through creation of experimental probabilistic thunderstorm forecasts and subsequent discussions between severe weather, aviation, and QPF forecast desks.
- Provide focused feedback to model developers on the performance of the experimental SSEF and deterministic models during severe thunderstorm episodes.

B. Aviation Thunderstorm-Impacts Component (Leader: AWC)

- Initial exploration of high-resolution convection-allowing models (CAMs) and SSEF to provide useful guidance to aviation forecasters for the creation of experimental probabilistic thunderstorm forecasts. These will focus on improving forecasts of timing, location, coverage (porosity), and tops of thunderstorms that are critical for the efficient management of the NAS.
- Assess the perceived value of probabilistic output and other unique fields from the SSEF, CAMs, CoSPA (after June 1), LAMP, etc. as guidance for creation of experimental probabilistic aviation thunderstorm products.
- Explore the utility of GOES-R Proving Ground products as part of an integrated data suite to support aviation forecaster decision-making.
- Test and evaluate the ability of aviation forecasters to create reliable probabilistic snapshot products for the occurrence of 40 dBZ echoes, especially in lines or clusters, and the likelihood of echo tops exceeding critical flight levels. The snapshot times of 21, 23, and 01z will be valid during the peak convective diurnal period.
- Explore creation of Day Two strategic forecasts for potential thunderstorm impacts on aviation using SREF probabilistic guidance, CAM output, etc.
- Familiarize air traffic management specialists from ATCSCC and airlines with experimental probabilistic thunderstorm forecasts. Obtain feedback from specialists to identify ways to improve forecast product display and effectively interpret probabilistic forecast information in order to more effectively manage NAS.
- Build cross cutting relationships between members of the severe weather, aviation, and QPF communities to strengthen collaborations focused on shared thunderstorm forecast challenges.
- Explore forecast consistency through creation of experimental probabilistic thunderstorm forecasts and subsequent discussions between severe weather, aviation, and QPF forecast desks.

- Expose the storm scale numerical modeling community to the challenges and requirements of the aviation community and explore ideas for common solutions.
- Share knowledge and experience between the Hazardous Weather Testbed and the fledgling Aviation Weather Testbed which will subsequently serve as a focus for aviation collaborations for years to come.

C. QPF Component (Leader: HPC)

- Explore utility of 00z high-resolution convection-allowing models (CAMs) and SSEF system to provide useful guidance to hydrometeorological forecasters in creation of experimental probabilistic 6-hr QPF products. Document strengths and limitations of high resolution models for precipitation forecasting, and determine appropriate ways to use operational mesoscale (e.g, NAM, GFS, SREF) and experimental CAMs/SSEF in a complementary manner.
- Assess the perceived value of probabilistic output and other unique guidance from the SSEF to aid forecasters in the formulation of experimental 6-hr QPF products, consistent with current HPC operational requirements.
- Using experimental and operational data sets, determine if forecasters can create reliable probabilistic QPF products for 0.5” and 1.0” precipitation thresholds, valid for the 18-00z and 00-06z periods encompassing the diurnal convective cycle.
- Collaborate with the DTC to develop and test traditional and new objective verification measures to assess the accuracy of the experimental probabilistic QPF forecasts and the SSEF guidance.
- Provide focused feedback to model developers on performance and model utility in a simulated forecasting environment, share ideas on innovative ways to extract information from CAMs and SSEF, and offer recommendations to operational modeling community on near-term and long-term model development needs to support improvements in QPF.
- Build cross cutting relationships between members of the severe weather, aviation, and QPF communities to strengthen collaborations focused on shared thunderstorm forecast challenges.
- Explore forecast consistency through creation of experimental probabilistic thunderstorm forecasts and subsequent discussions between severe weather, aviation, and QPF forecast desks.

- Share knowledge and experience between the Hazardous Weather Testbed and the Hydrometeorological Testbed–HPC, which serves as a focus for QPF collaborations.

9. Spring Experiment Web Site

A full description of all program objectives, types of model output, forecast products, evaluation and verification forms, a data archive, and other related links are available at the Spring Experiment web site:

http://hwt.nssl.noaa.gov/Spring_2010/

This web site is intended to support real time activities as well as additional research and reference after the conclusion of the program.

10. Dates and Location of the Spring Experiment

The 2010 Spring Experiment will run Monday- Friday 7:30 am – 4:00 pm from May 17 through June 18, 2010. The final Friday session on June 18 will end by noon as no forecast activities will take place on that day. **On each Monday, a brief orientation session will start the day to introduce participants to the HWT and the planned experimental activities in each of the three components.** Beginning May 17, a full range of in-house and external participants will staff the program. Full time participants will work for periods of one week, with part-time visiting scientists and forecasters participating on a 2-3 day basis (schedule permitting). Program operations will be conducted in the Hazardous Weather Testbed facility (Room 2380) located on the second floor of the NWC between the SPC and WFO Norman operations areas. Each full time weekly team will complete daily experimental forecasts and participate in evaluation and verification activities; part-time visitors can participate in daily activities at levels appropriate with their interest and expertise. Staffing typically will include SPC, AWC, and HPC forecasters, several SPC and NSSL scientists, and a number of visiting scientists, model developers, forecasters, university faculty, and graduate students. A list of weekly participants is found in Appendix A.

11. Daily Operations Schedule

Participants in the experiment will create experimental forecast products and conduct evaluation activities in the HWT from 7:30 am – 4:00 pm on Monday-Friday. Each afternoon at 3:00 pm a daily summary session is held, and each Friday a weekly summary will be conducted. We anticipate that many weekly participants will rotate through the activities in each component (severe, aviation, QPF) during the week, spending 1-2 days in each section. This will allow participants to experience a broad range of convective storm impacts and forecasting challenges, and gain a greater appreciation of the challenges faced by operational forecasters and those tasked with creating improved forecast guidance tools.

Participants are expected to perform forecast and evaluation activities in a collaborative manner, such that results reflect a consensus decision. A break for lunch is scheduled during the ~Noon-12:30 pm period, but may eat lunch while conducting program activities or at

their discretion any time during the day. Visitors may purchase lunch at a food court located on the south side of the first floor of the NWC. Below is a basic outline of the daily schedule for activities during the experiment; a more detailed description is found in Appendix B.

A. Severe Convective Storms Component

Daily activities conducted in northeast corner of HWT

Italics denotes Monday-only activities

7:30 am-8:00 am: Weekly Orientation. (Some morning forecast and evaluation activities will be truncated on Mondays to permit sufficient time for the orientation.)

7:30-8:15 am: Subjective verification of yesterday's experimental severe weather forecasts compared to severe storm report maps and post-processed "practically perfect" hindcasts.

8:15-10:30 am: In a semi-operational forecasting environment, the severe weather team will use guidance from 00z high resolution WRF and SSEF, and 09z SREF/12z operational models and observational data to formulate probabilistic severe storm forecasts valid for the 20-00z and 00-04z time periods. The forecasts will be made over a movable mesoscale domain placed over the part of the central-eastern US where the severe threat is deemed to be greatest and/or substantial forecasting challenges exist. The process will include collaboration discussions between the severe, aviation, and QPF components prior to product completion to enhance consistency among the convective forecasts.

10:30 am-noon: Subjective/objective evaluation of previous day's model guidance compared to observed radar and severe weather reports, focusing on the ability of the models to provide useful guidance to severe weather forecasters.

Noon-12:30 pm: Lunch.

12:30-2:30 pm: Update two-period severe weather forecasts focusing on use and perceived value of hourly guidance from the HRRR model. CoSPA forecasts are expected to be available in June and guidance from that system may also be examined for the afternoon update forecasts. The process will include collaboration discussions between the severe and aviation components prior to product completion to enhance consistency among the convective forecasts.

2:30-3:00 pm: Break and preparation for briefing.

3:00-4:00 pm: Daily briefing and discussion of today's forecast and evaluation activities from the severe weather, aviation, and QPF teams, summarizing new insights, preliminary findings, lessons learned, and topic areas needing further examination. On Fridays, a weekly wrap-up is provided.

B. Aviation Impacts Component

Daily activities conducted in northwest corner of HWT

Italics denotes Monday-only activities

7:30 am-8:00 am: Weekly Orientation. (Some morning forecast and evaluation activities will be truncated on Mondays to permit sufficient time for the orientation.)

7:30-8:15 am: Subjective verification of yesterday's experimental aviation thunderstorm forecasts compared to observed radar data and NAS flight disruption data.

8:15-10:30 am: In a semi-operational forecasting environment, the aviation desk will use guidance from 00z CAMs and SSEF, 09z SREF/12z operational model guidance, and morning HRRR, LAMP, and, after June 1, CoSPA guidance to create probabilistic aviation thunderstorm forecasts of areas with ≥ 40 dBZ echoes, including delineation of broken or solid lines, at 21, 23, and 01z. In addition, a probabilistic forecast of echo tops ≥ 250 will be made for only the 23z snapshot time. The forecasts will generally be made over a fixed domain covering parts of the central and eastern US, although this can be adjusted to focus on smaller regions when necessary. The process will include collaboration discussions between the severe, aviation, and QPF components prior to product completion to enhance consistency among the convective forecasts.

10:30 am-noon: Subjective/objective evaluation of previous day's model guidance compared to observed radar (and NAS flight disruption data), focusing on the ability of various guidance tools to provide useful guidance to aviation forecasters.

Noon-12:30 pm: Lunch.

12:30-2:30 pm:

Break into two aviation teams

Team 1: Morning Update Team

Update the morning snapshot forecasts using observational data and guidance from the HRRR, LAMP, CoSPA (after June 1), etc. The process will include collaboration discussions between the severe and aviation components prior to product completion to enhance consistency among the convective forecasts.

Team 2: Day 2 Strategic Forecast Team

Using latest SREF output, operational model guidance, and 12z CAM output, issue an experimental probabilistic forecast of ≥ 40 dBZ echoes over the same central-eastern US fixed domain valid for the Day 2 period of 18-00z. (This may be modified to conform to Collaborative Strategic Planning Process requirements.)

2:30-3:00 pm: Break and preparation for briefing.

3:00-4:00 pm: Daily briefing and discussion of today's forecast and evaluation activities from the severe weather, aviation, and QPF teams, summarizing new insights,

preliminary findings, lessons learned, and topic areas needing further examination. On Fridays, a weekly wrap-up is provided.

C. QPF Component

Daily activities conducted in north center part of HWT

Italics denotes Monday-only activities

7:30 am-8:00 am: Weekly Orientation. (Some morning forecast and evaluation activities will be truncated on Mondays to permit sufficient time for the orientation.)

7:30-8:15 am: Subjective verification of yesterday's experimental QPF products compared to NSSL QPE ("truth").

8:15-10:30 am: In a semi-operational forecasting environment, the QPF desk will use guidance from 00z CAMs and SSEF, 09z SREF/12z operational model guidance and observational data to create experimental probabilistic QPF products valid for 18-00z and 00-06z time periods. The forecasts will be over the same mesoscale domain selected for the HWT severe convective weather component, and will be for exceedance thresholds of 0.5" and 1.0" per 6 hrs. In addition, forecasts that contain a probabilistic 1" contour will include a maximum basin-average rainfall amount within the 1" region during the 6-hr period. The process will include collaboration discussions between the severe, aviation, and QPF components prior to product completion to enhance consistency among the convective forecasts.

10:30 am-noon: Subjective/objective evaluation of previous day's experimental model guidance compared to NSSL QPE, focusing on model and product ability to provide useful guidance to QPF forecasters.

Noon-12:30 pm: Lunch.

12:30-2:30 pm: QPF participants will work with either the aviation-impacts or severe weather teams during the afternoon forecast update activities.

2:30-3:00 pm: Break and preparation for briefing.

3:00-4:00 pm: Daily briefing and discussion of today's forecast and evaluation activities from the severe weather, aviation, and QPF teams, summarizing new insights, preliminary findings, lessons learned, and topic areas needing further examination. On Fridays, a weekly wrap-up is provided.

12. Acknowledgments

Special thanks and appreciation is extended to many people for their creative insights and assistance in Spring Experiment preparations, planning, and execution of numerous complex and ground-breaking technical and scientific activities. Without the combined efforts of many SPC and NSSL staff, the Spring Program could not be conducted. In particular, special thanks go to SPC's Gregg Grosshans for helping to establish model data flow and configuring the experimental forecasts for transmission and archival, and for helping to organize model display files, Andy Dean for data processing and graphics display support, and Jay Liang and Joe Byerly for expertise in configuring and upgrading hardware/software, network and workstations in the HWT. Expert NSSL technical support was provided by Brett Morrow, Jeff Horn, Steve Fletcher and Brad Sagowitz to address networking, data flow, hardware, and archive requirements. Linda Crank (SPC), Peggy Stogsdill (SPC), and Linda Foster (NSSL) ably assisted with logistical and budget support activities. J. J. Gourley's assistance in organizing the NSSL Seminar Series for visiting scientists during the experiment is also greatly appreciated.

The experimental activities could not take place without the dedicated collaborative efforts of many people at CAPS, EMC, GSD and NCAR who are working to enhance community efforts to improve severe weather forecasting. We acknowledge the expertise of CAPS scientists Ming Xue, Kelvin Droegemeier, Fanyou Kong, Kevin Thomas, Yunheng Wang, Jidong Gao, and Keith Brewster for outstanding efforts to develop and run the 26 member SSEF and the 1 km and 4 km WRF models; the National Institute for Computational Sciences/University of Tennessee, and the University of Oklahoma Supercomputing Center for Education and Research for providing technical support and computer facilities for the CAPS model runs; NCAR scientists Morris Weisman, Wei Wang, Greg Thompson, and Jimy Dudhia for developing the NCAR WRF runs and offering technical support and advice on the ARW system and configuration of the SSEF; EMC scientists Matt Pyle, Zavisla Janjic, Brad Ferrier, Jun Du, and Geoff DiMego for developing and contributing the WRF-NMM models and for scientific input and infrastructure support for the SSEF; Stan Benjamin, Steve Weygandt, John Brown, and Curtis Alexander at ESRL/GSD for development of the CONUS HRRR model and establishment of data flow to SPC for use in the Spring Experiment; Barb Brown, David Ahijevich, Jamie Wolff, Steve Sullivan, Paul Oldenberg and colleagues from the DTC for their expert development of the MET and MODE verification systems and special efforts to tailor it for application during the Spring Experiment; Bill McCaul, Jon Case, Scott Dembek and colleagues at NASA/SPoRT/Huntsville for implementation of the WRF Total Lightning Threat products in the NSSL-WRF model; Louie Grasso and Dan Lindsey at CIRA, and Jason Otkin and Justin Sieglaff at CIMSS for creating and providing simulated satellite imagery from the NSSL-WRF model; Wayne Feltz, Jason Brunner, and Justin Sieglaff of CIMSS for the enhanced-V/overshooting tops satellite products; John Mecikalski, Wayne MacKenzie, and John Walker of UAH, and Wayne Feltz and Justin Sieglaff of CIMSS and Kris Bedka of NASA/Langley for convective initiation satellite products; and CIMSS scientists Ralph Petersen and Bob Aune for access to their Nearcast model.

Special thanks go to the GOES-R Program Office, the National Weather Service, and the National Severe Storms Laboratory for providing visitor travel support. We further wish to recognize the full support of SPC and NSSL management; and the numerous contributions

and insights provided by the many participants who clearly demonstrated the value of collaborative experiments involving the research, academic, and forecasting communities, and whose presence and enthusiasm resulted in a positive learning experience for everyone.

Appendix A: Spring Experiment Participant Schedule (5/21/10)

Weekly Calendar of EFP Spring Experiment Participants – 2010

Mon May 17	Tue May 18	Wed May 19	Thu May 20	Fri May 21
Jonathan Garner Faye Barthold Katie Burtis Jason Levit Bryan Tilley Dave Nadler Tara Jensen Tom Hamill Bill Gallus Jon Case Kris Bedka Justin Sieglaff John Walker Eric Bruning Steve Weiss Chris Siewert Patrick Marsh Adam Clark	Jonathan Garner Faye Barthold Katie Burtis Jason Levit Bryan Tilley Dave Nadler Tara Jensen Tom Hamill Bill Gallus John Huhn Jon Case Kris Bedka Justin Sieglaff John Walker Eric Bruning Steve Weiss Chris Siewert Patrick Marsh Adam Clark Brian Kolts Glenn Robinson	Jonathan Garner Faye Barthold Katie Burtis Jason Levit Bryan Tilley Dave Nadler Tara Jensen Tom Hamill Bill Gallus John Huhn Jon Case Kris Bedka Justin Sieglaff John Walker Eric Bruning Steve Weiss Chris Siewert Patrick Marsh Adam Clark Brian Kolts	Jonathan Garner Faye Barthold Katie Burtis Jason Levit Bryan Tilley Dave Nadler Tara Jensen Tom Hamill Bill Gallus Jon Case Kris Bedka Justin Sieglaff John Walker Eric Bruning Steve Weiss Chris Siewert Patrick Marsh Adam Clark Brian Kolts	Faye Barthold Katie Burtis Jason Levit Bryan Tilley Dave Nadler Tara Jensen Tom Hamill Jon Case John Walker Eric Bruning Steve Weiss Chris Siewert Patrick Marsh Adam Clark
Mon May 24	Tue May 25	Wed May 26	Thu May 27	Fri May 28
Jonathan Garner Bob Oravec Bruce Entwistle Bob Hoenisch Chuck Jones Nelson Vaz Geoff Manikin Dave Ahijevych Kevin Manning Louie Grasso Becky Selin Wayne Feltz Jason Otkin Wayne MacKenzie Eric Bruning Bill Campbell Bonnie Reed Steve Weiss Chris Siewert Patrick Marsh Adam Clark Greg Noonan	Jonathan Garner Bob Oravec Bruce Entwistle Bob Hoenisch Chuck Jones Nelson Vaz Geoff Manikin Dave Ahijevych Kevin Manning Louie Grasso Becky Selin Wayne Feltz Jason Otkin Wayne MacKenzie Eric Bruning Bill Campbell Bonnie Reed Steve Weiss Chris Siewert Patrick Marsh Adam Clark Brian Proctor Greg Noonan	Jonathan Garner Bob Oravec Bruce Entwistle Bob Hoenisch Chuck Jones Nelson Vaz Geoff Manikin Dave Ahijevych Kevin Manning Louie Grasso Becky Selin Wayne Feltz Jason Otkin Wayne MacKenzie Eric Bruning Bill Campbell Bonnie Reed Steve Weiss Chris Siewert Patrick Marsh Adam Clark Louis Uccellini Greg Mandt Don Berchoff Greg Noonan	Jon Racy Bob Oravec Bruce Entwistle Bob Hoenisch Chuck Jones Nelson Vaz Geoff Manikin Dave Ahijevych Kevin Manning Louie Grasso Becky Selin Wayne Feltz Jason Otkin Wayne MacKenzie Eric Bruning Bill Campbell Bonnie Reed Steve Weiss Chris Siewert Patrick Marsh Adam Clark Louis Uccellini Greg Mandt Don Berchoff Kevin Johnston Michael Elliot	Jon Racy Bob Oravec Bruce Entwistle Bob Hoenisch Chuck Jones Nelson Vaz Geoff Manikin Dave Ahijevych Kevin Manning Louie Grasso Becky Selin Jason Otkin Wayne MacKenzie Eric Bruning Bill Campbell Bonnie Reed Steve Weiss Chris Siewert Patrick Marsh Adam Clark Kevin Johnston Louis Uccellini Greg Mandt Don Berchoff

Mon May 31	Tue Jun 1	Wed Jun 2	Thu Jun 3	Fri Jun 4
Memorial Day Holiday – No HWT Operations	Jared Guyer Bruce Entwistle Steve Silberberg Dave Novak Tom Clemmons Mark O'Malley Jeff Hovis Gene Petrescu Glenn Lader Phil Shafer Matt Pyle Michelle Harrold Tara Jensen Curtis Alexander Evan Kuchera Bill McCaul Chris Rozoff Dan Hartung Steve Weiss Chris Siewert Patrick Marsh Adam Clark Brad Sherman	Jared Guyer Bruce Entwistle Steve Silberberg Dave Novak Tom Clemmons Mark O'Malley Jeff Hovis Gene Petrescu Glenn Lader Phil Shafer Matt Pyle Michelle Harrold Tara Jensen Curtis Alexander Evan Kuchera Bill McCaul Chris Rozoff Dan Hartung Steve Weiss Chris Siewert Patrick Marsh Adam Clark Tom Lloyd Brad Sherman	Jared Guyer Bruce Entwistle Steve Silberberg Dave Novak Tom Clemmons Mark O'Malley Jeff Hovis Gene Petrescu Glenn Lader Phil Shafer Matt Pyle Michelle Harrold Tara Jensen Curtis Alexander Evan Kuchera Bill McCaul Chris Rozoff Dan Hartung Steve Weiss Chris Siewert Patrick Marsh Adam Clark Tom Lloyd Brad Sherman	Jared Guyer Bruce Entwistle Steve Silberberg Dave Novak Tom Clemmons Mark O'Malley Jeff Hovis Gene Petrescu Glenn Lader Phil Shafer Matt Pyle Michelle Harrold Tara Jensen Curtis Alexander Evan Kuchera Bill McCaul Chris Rozoff Dan Hartung Steve Weiss Chris Siewert Patrick Marsh Adam Clark Brad Sherman
Mon Jun 7	Tue Jun 8	Wed Jun 9	Thu Jun 10	Fri Jun 11
Jeremy Grams Bruce Sullivan Steve Burbak Dan Vallee Steve Hrebenach Mark Burger Bill Ward Jess Charba Doug Streu Matt Lorentson Jamie Wolff Lance Bosart Tom Galarneau Russ Schumacher H. Iskenderian Lee Cronce Geoff Stano Sarah Monette Steve Weiss Chris Siewert Patrick Marsh Adam Clark Aaron Johnson	Jeremy Grams Bruce Sullivan Steve Burbak Dan Vallee Steve Hrebenach Mark Burger Bill Ward Jess Charba Doug Streu Matt Lorentson Jamie Wolff Lance Bosart Tom Galarneau Russ Schumacher H. Iskenderian Lee Cronce Geoff Stano Sarah Monette Steve Weiss Chris Siewert Patrick Marsh Adam Clark Aaron Johnson	Jeremy Grams Bruce Sullivan Steve Burbak Dan Vallee Steve Hrebenach Mark Burger Bill Ward Jess Charba Doug Streu Matt Lorentson Jamie Wolff Lance Bosart Tom Galarneau Russ Schumacher H. Iskenderian Lee Cronce Geoff Stano Sarah Monette Steve Weiss Chris Siewert Patrick Marsh Adam Clark Aaron Johnson	Jeremy Grams Bruce Sullivan Steve Burbak Dan Vallee Steve Hrebenach Mark Burger Bill Ward Jess Charba Doug Streu Jamie Wolff Lance Bosart Tom Galarneau Russ Schumacher H. Iskenderian Lee Cronce Geoff Stano Sarah Monette Steve Weiss Chris Siewert Patrick Marsh Adam Clark Aaron Johnson	Bruce Sullivan Steve Burbak Dan Vallee Steve Hrebenach Mark Burger Bill Ward Jess Charba Doug Streu Jamie Wolff Lance Bosart Tom Galarneau Russ Schumacher H. Iskenderian Lee Cronce Geoff Stano Sarah Monette Steve Weiss Chris Siewert Patrick Marsh Adam Clark Aaron Johnson

Mon Jun 14	Tue Jun 15	Wed Jun 16	Thu Jun 17	Fri Jun 18
Ryan Jewell	Ryan Jewell	Ryan Jewell	Ryan Jewell	Rich Bann
Rich Bann	Rich Bann	Rich Bann	Rich Bann	K. Eagle-Brogan
K. Eagle-Brogan	K. Eagle-Brogan	K. Eagle-Brogan	K. Eagle-Brogan	Zavisa Janjic
Zavisa Janjic	Zavisa Janjic	Zavisa Janjic	Zavisa Janjic	Paul Lee
Paul Lee	Paul Lee	Paul Lee	Paul Lee	Gail Hartfield
Gail Hartfield	Gail Hartfield	Gail Hartfield	Gail Hartfield	Bobby Prentice
Bobby Prentice	Bobby Prentice	Bobby Prentice	Bobby Prentice	Tony Eckel
Tony Eckel	Tony Eckel	Tony Eckel	Tony Eckel	Isidora Jankov
Isidora Jankov	Isidora Jankov	Isidora Jankov	Isidora Jankov	John Brown
John Brown	John Brown	John Brown	John Brown	Ryan Torn
Ryan Torn	Ryan Torn	Ryan Torn	Ryan Torn	Ralph Petersen
Ralph Petersen	Ralph Petersen	Ralph Petersen	Ralph Petersen	Bob Aune
Bob Aune	Bob Aune	Bob Aune	Bob Aune	Jordan Gerth
Jordan Gerth	Jordan Gerth	Jordan Gerth	Jordan Gerth	Jim Gurka
Jim Gurka	Jim Gurka	Jim Gurka	Jim Gurka	Steve Weiss
Dan Lindsey	Dan Lindsey	Dan Lindsey	Dan Lindsey	Chris Siewert
Steve Weiss	Steve Weiss	Steve Weiss	Steve Weiss	Patrick Marsh
Chris Siewert	Chris Siewert	Chris Siewert	Chris Siewert	Adam Clark
Patrick Marsh	Patrick Marsh	Patrick Marsh	Patrick Marsh	
Adam Clark	Adam Clark	Adam Clark	Adam Clark	

EFP Spring Experiment 2010 Participants and Affiliations (5/21/10)

Primarily EWP * GOES-R Observer + Part-Time Visitor (1-3 days)

Week of May 17

Jonathan Garner (NOAA/NWS/NCEP SPC)
Steve Weiss (NOAA/NWS/NCEP SPC)
Chris Siewert (CIMMS/SPC GOES-R Proving Ground, Norman, OK)
Patrick Marsh (NOAA/OAR/NSSL)
Adam Clark (NOAA/OAR/NSSL)
Jason Levit (NOAA/NWS/NCEP/AWC, Kansas City, MO)
Katie Burtis (NOAA/NWS/NCEP/AWC, Kansas City, MO)
Faye Barthold (NOAA/NWS/NCEP/HPC, Camp Springs, MD)
Tara Jensen (NCAR/Developmental Testbed Center, Boulder, CO)
Tom Hamill (NOAA/OAR/ESRL/PSD Boulder CO)
Dave Nadler (NOAA/NWS Huntsville, AL)
Bryan Tilley (NOAA/NWS White Lake-Detroit, MI)
+Glenn Robinson (Environment Canada, Toronto)
Bill Gallus (Iowa State University, Ames, IA)
+John Huhn (Mitre Corp/FAA, Herndon, VA)
Jon Case (ENSCO Inc./SPoRT Center, Huntsville, AL)
#Justin Sieglaff (CIMSS/UW, Madison, WI)
#Kris Bedka (SSAI/NASA Langley Research Center, Hampton, VA)
John Walker (UAH, Huntsville, AL)
#Eric Bruning (CICS/University of Maryland, College Park, MD)
+Brian Kolts (FirstEnergy, Akron, OH)

Week of May 24

Jonathan Garner (NOAA/NWS/NCEP SPC)
Jon Racy (NOAA/NWS/NCEP SPC)
Steve Weiss (NOAA/NWS/NCEP SPC)
Chris Siewert (CIMMS/SPC GOES-R Proving Ground, Norman, OK)
Patrick Marsh (NOAA/OAR/NSSL)
Adam Clark (NOAA/OAR/NSSL)
Bruce Entwistle (NOAA/NWS/NCEP/AWC, Kansas City, MO)
Bob Oravec (NOAA/NWS/NCEP/HPC, Camp Springs, MD)
+Louis Uccellini (NOAA/NWS/NCEP Camp Springs, MD)
+Greg Mandt (NOAA/NESDIS, Greenbelt, MD)
+Don Berchhoff (NOAA/NWS/OST Silver Spring, MD)
+Kevin Johnston (FAA Command Center, Herndon, VA)
David Ahijevych (NCAR/Developmental Testbed Center, Boulder, CO)
Geoff Manikin (NOAA/NCEP/EMC Camp Springs, MD)
Bob Hoenisch (NOAA/NWS Great Falls, MT)
Nelson Vaz (NOAA/NWS Upton, NY)
Chuck Jones (NOAA/NWS Albuquerque, NM)
Kevin Manning (NCAR, Boulder, CO)
Louie Grasso (CIRA/CSU, Fort Collins, CO)
Rebecca Selin (CSU/AFWA, Bellevue, NE)
#Wayne Feltz (CIMSS/UW, Madison, WI)
Jason Otkin (CIMSS-UW, Madison, WI)
Wayne Mackenzie (UAH, Huntsville, AL)
#Eric Bruning (CICS/University of Maryland, College Park, MD)
*Bill Campbell (NOAA/NWS/OST, Silver Spring, MD)
*Bonnie Reed (NOAA/NWS/OST, Silver Spring, MD)
+Brian Proctor (Environment Canada, Edmonton)
+Michael Elliot (Environment Canada, Montreal)
+Greg Noonan (NOAA/NWS/CRH Kansas City, MO)

Week of June 1

Jared Guyer (NOAA/NWS/NCEP SPC)
Steve Weiss (NOAA/NWS/NCEP SPC)
Chris Siewert (CIMMS/SPC GOES-R Proving Ground, Norman, OK)
Patrick Marsh (NOAA/OAR/NSSL)
Adam Clark (NOAA/OAR/NSSL)
Bruce Entwistle (NOAA/NWS/NCEP/AWC, Kansas City, MO)
Steve Silberberg (NOAA/NWS/NCEP/AWC, Kansas City, MO)
Dave Novak (NOAA/NWS/NCEP HPC, Camp Springs, MD)
Michelle Harrold (NCAR/Developmental Testbed Center, Boulder, CO)
+Tara Jensen (NCAR/Developmental Testbed Center, Boulder, CO)
Matt Pyle (NOAA/NWS/NCEP/EMC, Camp Springs, MD)
Gene Petrescu (NOAA/NWS Anchorage, AK)
Tom Clemmons (NOAA/NWS Flagstaff, AZ)
Glenn Lader (NOAA/NWS Tucson, AZ)
Jeff Hovis (NOAA/NWS Charleston, WV)
Mark O'Malley (NOAA/NWS Pleasant Hill-Kansas City, MO)
Phil Shafer (NOAA/NWS/MDL, Silver Spring, MD)
Evan Kuchera (AFWA, Bellevue, NE)
Curtis Alexander (NOAA/OAR/ESRL/GSD Boulder, CO)
Bill McCaul (USRA/NASA/SPoRT, Huntsville, AL)
Chris Rozoff (CIMSS/UW, Madison, WI)
Dan Hartung (CIMSS/UW, Madison, WI)
Brad Sherman (FAA/Command Center, Herndon, VA)

Week of June 7

Jeremy Grams (NOAA/NWS/NCEP SPC)
Steve Weiss (NOAA/NWS/NCEP SPC)
Chris Siewert (CIMMS/SPC GOES-R Proving Ground, Norman, OK)
Patrick Marsh (NOAA/OAR/NSSL)
Adam Clark (NOAA/OAR/NSSL)
Steve Burbak (NOAA/NWS/NCEP/AWC, Kansas City, MO)
Bruce Sullivan (NOAA/NWS/NCEP/HPC, Camp Springs, MD)
Jamie Wolff (NCAR/Developmental Testbed Center, Boulder, CO)
Dan Vallee (NOAA/NWS Pocatello, ID)
Stephen Hrebenach (NOAA/NWS Wilmington, OH)
Mark Burger (NOAA/NWS Eureka, CA)
Bill Ward (NOAA/NWS Honolulu, HI)
+Jess Charba (NOAA/NWS/MDL, Silver Spring, MD)
+Matt Lorentson (NOAA/NWS/FAA Command Center, Herndon, VA)
+Bill Kuo (NCAR/Developmental Testbed Center, Boulder, CO)
Doug Streu (NOAA/NWS/FAA Oklahoma City, OK)
Lance Bosart (University at Albany/SUNY, Albany, NY)
Tom Galarneau (University at Albany/SUNY, Albany, NY)
Russ Schumacher (Texas A&M University, College Station, TX)
Haig Iskenderian (Lincoln Lab/MIT, Lexington, MA)
#Lee Cronce (CIMSS/UW, Madison, WI)
#Sarah Monette (CIMSS/UW, Madison, WI)
#Geoff Stano (ENSCO Inc./SPoRT Center, Huntsville, AL)
Aaron Johnson (University of Oklahoma, Norman, OK)

Week of June 14

Ryan Jewell (NOAA/NWS/NCEP SPC)
Steve Weiss (NOAA/NWS/NCEP SPC)
Chris Siewert (CIMMS/SPC GOES-R Proving Ground, Norman, OK)
Patrick Marsh (NOAA/OAR/NSSL)
Adam Clark (NOAA/OAR/NSSL)
Karen Eagle-Brogan (NOAA/NWS/NCEP/AWC, Kansas City, MO)
Rich Bann (NOAA/NWS/NCEP/HPC, Camp Springs, MD)
Paul Lee (NOAA/NWS/NCEP/OPC, Camp Springs, MD)
Zavisa Janjic (NOAA/NWS/NCEP/EMC, Camp Springs, MD)
Isidora Jankov (NCAR/Developmental Testbed Center and NOAA/OAR/ESRL/GSD, Boulder, CO)
+Tara Jensen (NCAR/Developmental Testbed Center, Boulder, CO)
+Paul Oldenburg (NCAR/Developmental Testbed Center, Boulder, CO)
Tony Eckel (NOAA/NWS/OST, Seattle, WA)
Gail Hartfield (NOAA/NWS Raleigh, NC)
Bob Prentice (NOAA/NWS/FAA Oklahoma City, OK)
John Brown (NOAA/OAR/ESRL/GSD Boulder, CO)
*Jim Gurka (NOAA/NESDIS, Greenbelt, MD)
Ryan Torn (University at Albany/SUNY, Albany, NY)
John Huhn (Mitre Corp/FAA, Herndon, VA)
+Tom Lloyd (JetBlue, New York, NY)
Bob Aune (CIMSS/UW, Madison, WI)
Ralph Petersen (CIMSS/UW, Madison, WI)
Jordan Gerth (CIMSS/UW, Madison, WI)
Dan Lindsey (CIRA/CSU, Fort Collins, CO)

Appendix B: Instructions for Creating and Submitting Experimental Severe Thunderstorm Forecasts

1. Experimental Severe Thunderstorm Forecast Graphics

Preliminary two-period severe weather forecasts will be issued in the morning and **final** two-period forecasts will be issued in the afternoon. The forecast valid time periods will be 20-00z and 00-04z. The severe weather forecast graphics will be similar in format to operational SPC outlooks, except only total severe storm probability contours will be formulated (no categorical outlook, and no general thunderstorms will be forecast). The same probability contours used in the operational outlooks will be used for the severe forecasts (5, 15, 30, 45, and 60 %); an area delineating potential for significant severe storms will be included when the probability for significant severe is 10% or greater. The Probability-to-Categorical conversion for total severe is identical to that used for the SPC Day 2 Outlook, and is shown below.

2. Drawing and Saving the Experimental Forecasts in NMAP

a. For the preliminary and final forecasts, the forecaster will draw in NMAP separate probability contours for each valid period, and will save each forecast as a separate graphic product. The process will utilize NMAP software that is used in SPC operations. When saving each experimental forecast graphic, the following modifications are required:

- 1) In the format outlook box, *change valid time to 2000z to 0000z (or 0000z to 0400z)*
- 2) In the product save box, *replace "outlook" with "svr_prelim" or "svr_final"*

b. Enter command in xterm window: sp10bg STN svr forecast # (such as sp10bg OKC svr final 2)

STN is METAR centerpoint ID, **forecast** is "prelim" or "final", and # is NAWIPS workstation number (1-6) where the graphic is created. This script archives the severe weather forecast, attaches date/time to the graphics file, and sends graphics to the web page.

3. Completing Model Discussion Section on Internal Web Page

- a. On HWT Spring Experiment web page click on Experimental Forecast Generation (Severe)
- b. Click on "Preliminary" or "Final" and the two-period forecast graphics will appear
- c. Complete Discussion Text Box and when finalized, click on Submit.

! Probability to Categorical Outlook Conversion	
(SIGNIFICANT SEVERE area needed where denoted by hatching - otherwise default to next lower category)	
Outlook Probability	Combined TORN, WIND, and HAIL
5%	SEE TEXT
15%	SLGT
30%	SLGT
45%	MDT
60%	HIGH

Appendix C: Instructions for Creating and Submitting Experimental Aviation Thunderstorm Forecasts

Day 1 Forecasts

1. Experimental Day 1 Aviation Thunderstorm Forecast Graphics

Preliminary probabilistic thunderstorm (reflectivity ≥ 40 dBZ) forecasts for three snapshot times at 21, 23, and 01z will be issued in the morning, and **final** snapshot forecasts for the same valid times will be updated in the afternoon. The graphics will delineate the probability of ≥ 40 dBZ echoes using categorical descriptive contours of Slight, Moderate, and High, where Slight=25%, Moderate=50%, and High=75%. If a broken or solid thunderstorm line is predicted, it will be indicated on the snapshot graphic. In addition, a separate graphic containing the probability of thunderstorm tops \geq FL250 (25,000 feet) will be produced for the 23z forecast time. The same descriptive contours of Slight, Moderate, and High will be used for the thunderstorm top forecast.

2. Drawing and Saving the Experimental Day 1 Forecasts in NMAP

a. For the preliminary and final forecasts, the forecaster will draw in NMAP separate graphics containing categorical contours for each valid time, and will save each forecast time as a separate graphic product. The process will utilize NMAP software that is used in SPC operations. When saving each experimental 40dBZ reflectivity forecast graphic, the following modifications are required:

- 1) In the format outlook box, *change valid time to 2100z to 2100z (or 2300z to 2300z; or 0100z to 0100z)*
- 2) In the product save box, *replace “outlook” with “avn_tstm_prelim” or “avn_tstm_final”*
- 3) Enter command in xterm window: sp10bg STN avntstm forecast # (such as sp10bg OKC avn final 2)

When saving the experimental echo top ≥ 250 forecast valid at 23z, the following modifications are required:

- 4) In the format outlook box, *change valid time to 2300z to 2300z*
- 5) In the product save box, *replace “outlook” with “avn_top_prelim” or “avn_top_final”*
- 6) Enter command in xterm window: sp10bg STN avntop forecast # (such as sp10bg OKC avn final 2)

STN is METAR centerpoint ID, **forecast** is “prelim” or “final”, and # is NAWIPS workstation number (1-6) where the graphic is created. This script archives the aviation forecast, attaches date/time to the graphics file, and sends graphics to the web page.

3. Completing the Day 1 Model Discussion Section on Internal Web Page

- a. On HWT Spring Experiment web page click on Experimental Forecast Generation (Aviation)
- b. Click on “Day 1” and the forecast graphics for each snapshot time will appear
- c. Complete Discussion Text Box and when finalized, click on Submit.

Day 2 Forecast

4. Experimental Day 2 Aviation Forecast Graphics

A probabilistic thunderstorm (reflectivity ≥ 40 dBZ) forecast valid for the 18-00z Day 2 period will be issued in the afternoon. The graphics will delineate the probability of ≥ 40 dBZ echoes using categorical descriptive contours of Slight, Moderate, and High, where Slight=25%, Moderate=50%, and High=75%.

2. Drawing and Saving the Experimental Day 2 Forecast in NMAP

a. The forecaster will draw and save a single graphic in NMAP containing categorical contours for the 18-00z Day 2 period. The process will utilize NMAP software that is used in SPC operations. When saving the Day 2 experimental 40dBZ reflectivity forecast graphic, the following modifications are required:

- 1) In the format outlook box, *change valid dates to the correct Day2 and Day3 dates, and change valid time to 1800z to 0000z*
- 2) In the product save box, *replace "outlook" with "avn_tstm"*
- 3) Enter command in xterm window: sp10bg STN avn day2 # (such as sp10bg OKC avn day2 2)
STN is METAR centerpoint ID and # is NAWIPS workstation number (1-6) where the graphic is created. This script archives the day 2 thunderstorm forecast, attaches date/time to the graphics file, and sends graphics to the web page.

3. Completing the Day 2 Model Discussion Section on Internal Web Page

- a. On HWT Spring Experiment web page click on Experimental Forecast Generation (Aviation)
- b. Click on "Day 2" and the forecast graphics for each snapshot time will appear
- c. Complete Discussion Text Box and when finalized, click on Submit.

Appendix D: Instructions for Creating and Submitting Experimental QPF Products

1. Experimental QPF Graphics

Two-period QPF products for the probability of exceeding (POE) 0.5 inch/6 hrs and the probability of exceeding 1.0 inch/6 hrs will be issued in the morning. The forecast valid time periods will be 18-00z and 00-06z. The probabilistic QPF graphics are analogous to several operational HPC forecast products (such as excessive rainfall and heavy snow) in that categorical descriptive terms of Slight, Moderate, and High are used to denote forecast probabilities. For the experimental QPF products, Slight=25%, Moderate=50%, and High=75%. In addition to the categorical contours, a predicted maximum 6 hr rainfall amount within the highest categorical contour for the 1 inch POE will be included in the graphic. If no contours are drawn on the 1 inch POE graphic, a maximum amount is not necessary.

2. Drawing and Saving the Experimental Forecasts in NMAP

a. For the QPF forecasts, the forecaster will draw in NMAP separate categorical contours for both exceedance thresholds for the first valid period, and repeat the process for the second valid period. Each of the four graphics will be saved as a separate product. The process will utilize NMAP software that is used in SPC operations. When saving each experimental forecast graphic, the following modifications are required:

- 1) In the format outlook box, *change valid time to 1800z to 0000z (or 0000z to 0600z)*
- 2) In the product save box, *replace "outlook" with "qpf_50" or "qpf_100"*

b. Enter command in xterm window: sp10bg STN qpf final # (such as sp10bg OKC qpf final 2)

STN is METAR centerpoint ID, and # is NAWIPS workstation number (1-6) where the graphic is created. This script archives the QPF forecast, attaches date/time to the graphics file, and sends graphics to the web page.

3. Completing Model Discussion Section on Internal Web Page

- a. On HWT Spring Experiment web page click on Experimental Forecast Generation (QPF)
- b. Click on "QPF Products and the two-period forecast graphics will appear
- c. Complete Discussion Text Box and when finalized, click on Submit.

Appendix E: Practically Perfect Forecasts

(From Brooks, H. E., M. Kay, and J. A. Hart, 1998: Objective limits on forecasting skill of rare events. *Preprints*, 19th Conference on Severe Local Storms, Minneapolis, Minnesota, American Meteorological Society, 552-555.)

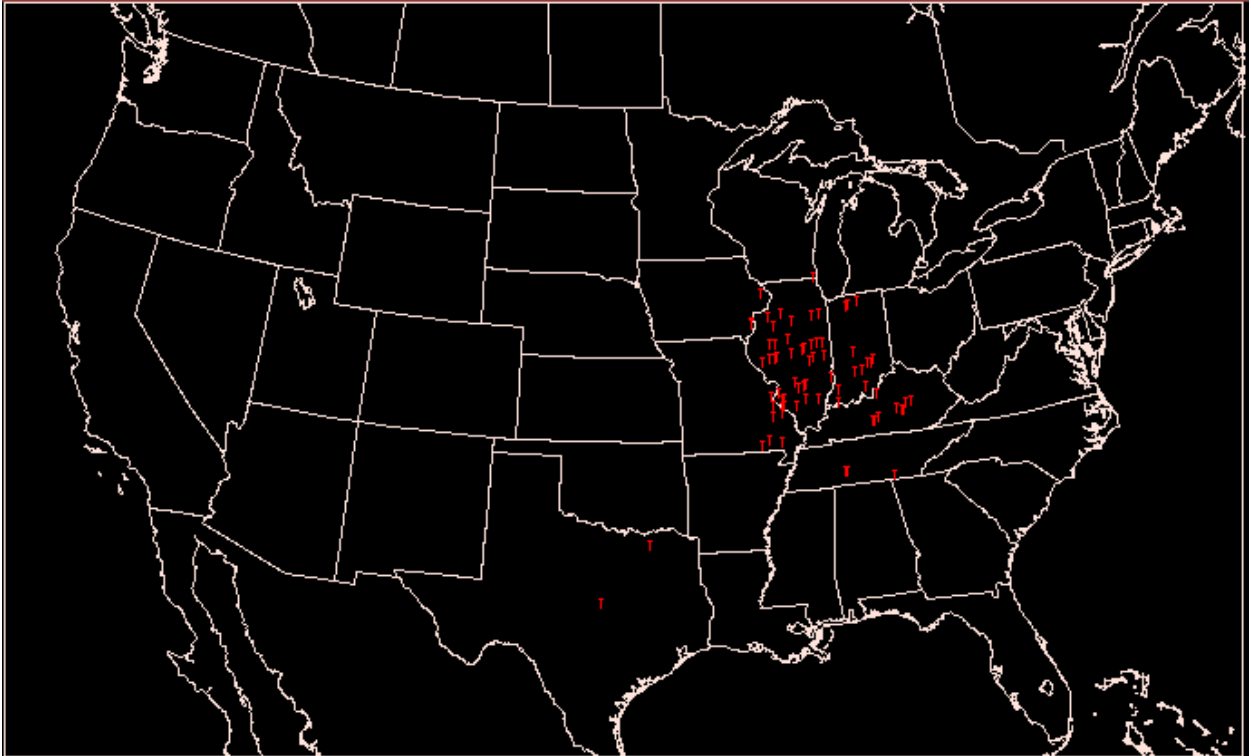
Severe weather forecasts such as SPC outlook and watch products are issued with the explicit expectation that there will be “false alarms” (parts of the forecast for which there are no events) and “missed detections” (events which are not included in the forecast). Thus, the expected range of values of the probability of detection (POD) or false alarm rate (FAR), for example, does not run from 0 to 1 in practice. The concept of a “practically” perfect (PP) forecast can then be used to estimate the minimum and maximum scores that a forecaster could reasonably be expected obtain given real world distributions of severe weather reports and the low predictability of specific severe convective storms in advance. In general, that range will be much smaller than the absolute minimum and maximum, but will provide a range over which meaningful forecast performance can be judged.

To compute the PP forecast, reports of severe weather are recorded on a grid with each grid box representing an area 80 x 80 km. (This grid corresponds to SPC Outlook products where probability values correspond to a probability within 25 miles of each grid point.) All severe weather reports are considered equal and the computation considers only whether a box has had an event or not. The PP forecast is then created by smoothing the events using nonparametric density estimation with a two dimensional Gaussian kernel. Specifically, at each grid point in the domain, the PP forecast value, f , is given by

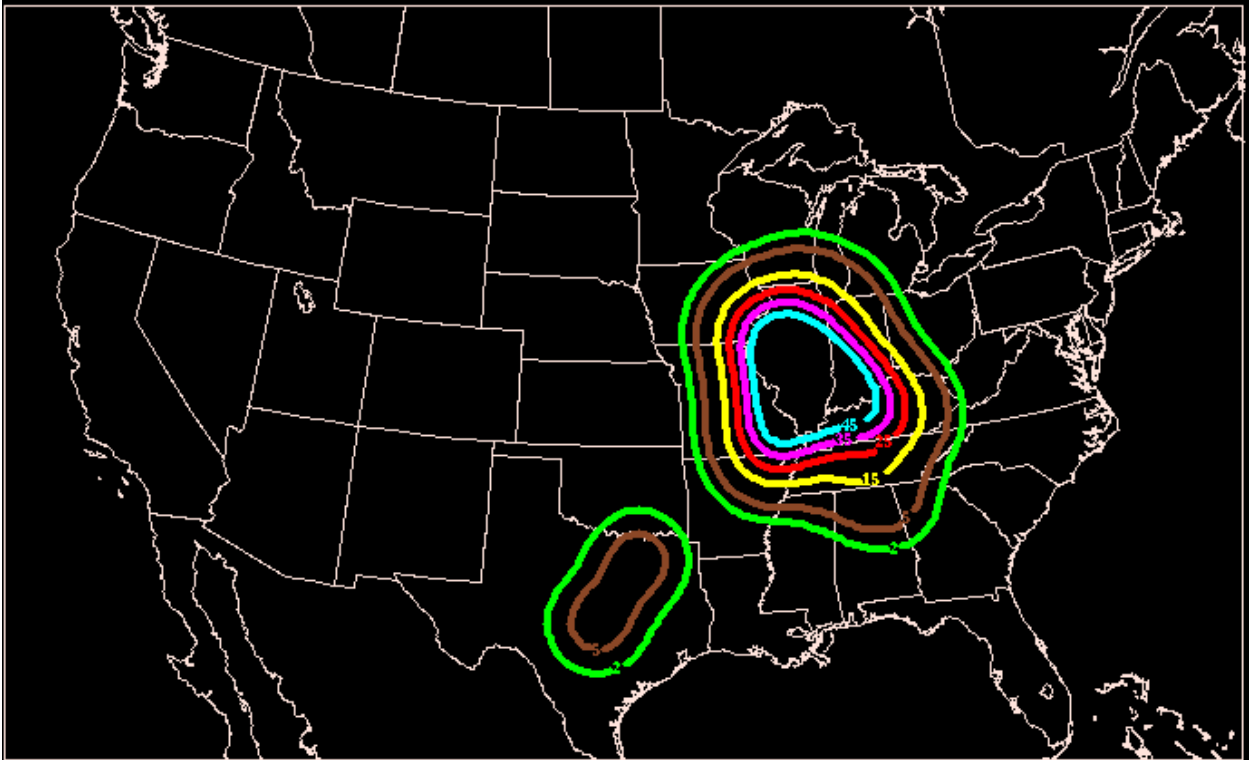
$$f = \sum_{n=1}^N \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left[\frac{d_n}{\sigma}\right]^2\right)$$

where d_n is the distance from the forecast grid point to the n -th location that had an event occur, N is the total number of grid points with events, and σ is a weighting function that can be interpreted as the confidence one has in the location of the forecast event. Increasing σ is equivalent to increasing the uncertainty associated with the forecast as one would do with increasing lead time of the forecast. That is, in the context of severe weather forecasting, very small σ can be thought of as being associated with the warning stage, while larger σ is associated with the watch or convective outlook stages. For SPC forecasts a value of 3 used.

Examples of practically perfect forecasts based on actual severe weather reports are shown on the next two pages.

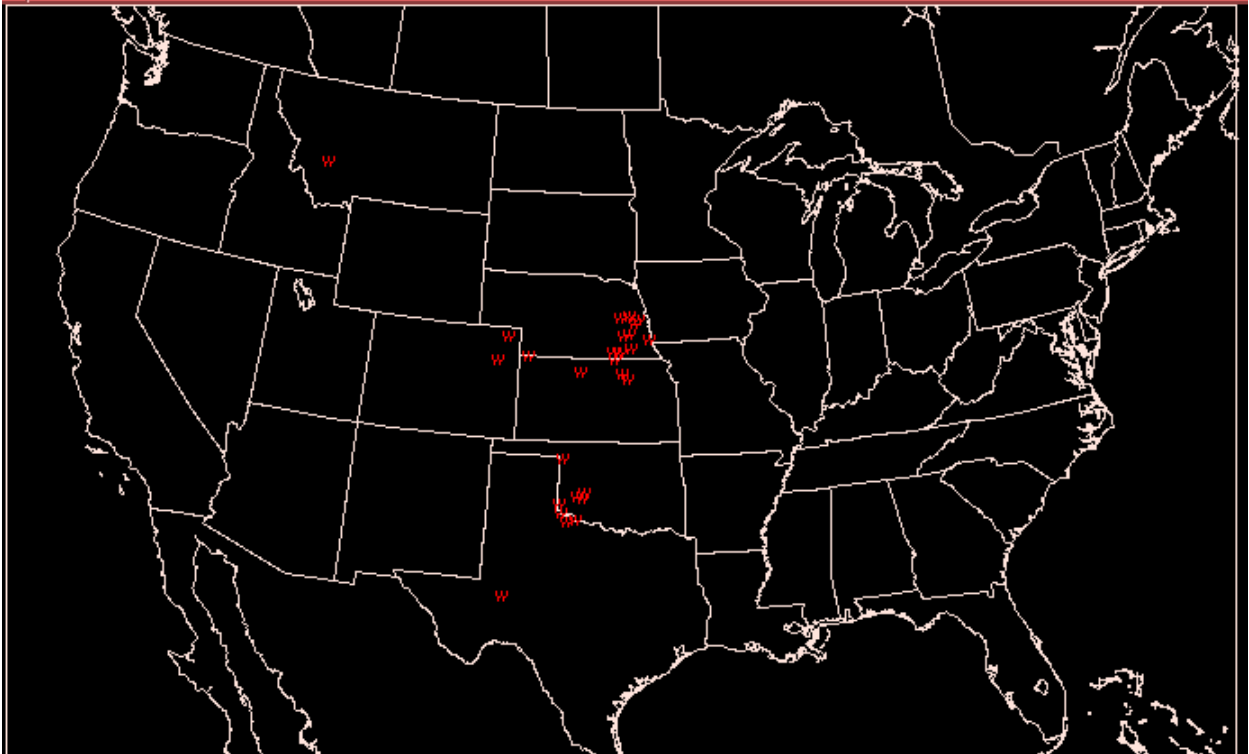


Tornado Distribution for 19960419

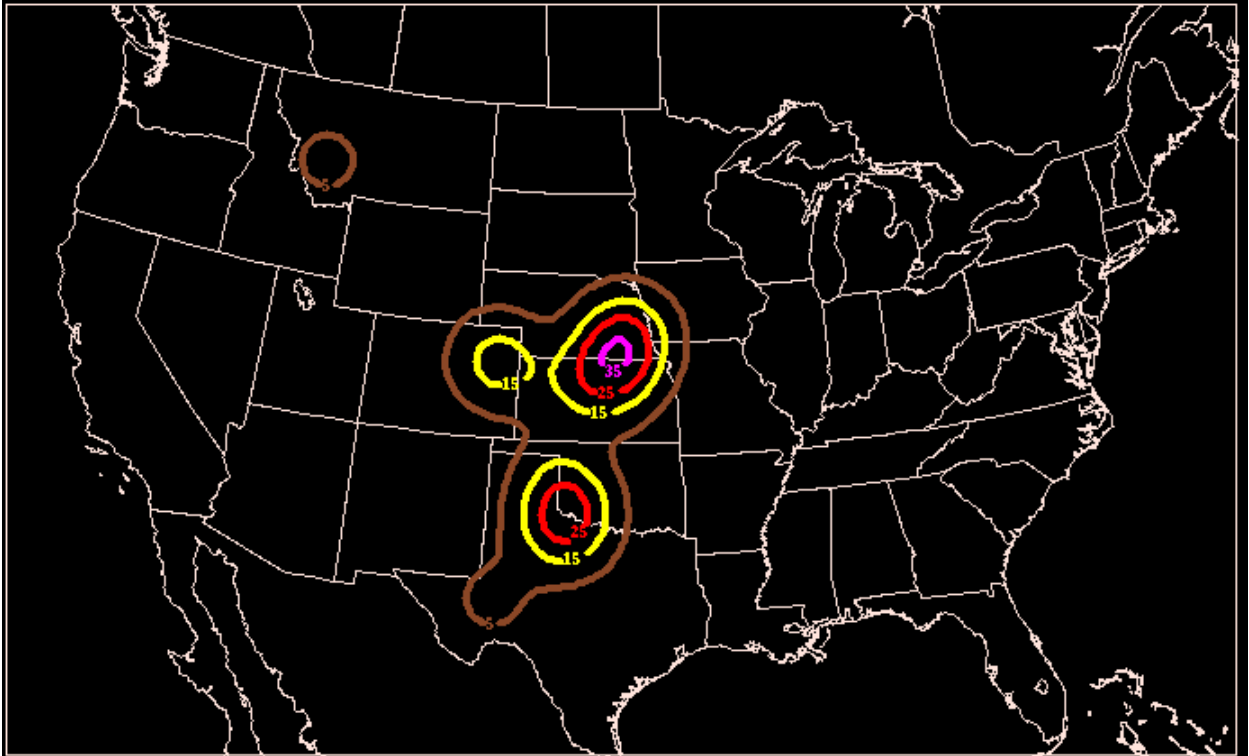


Practically-Perfect Hindcast

19 April 1996 tornado reports (top) and PP forecast (bottom) based on tornado reports



Wind Distribution for 19960522



Practically- Perfect Hindcast

22 May 1996 wind damage reports (top) and PP forecast (bottom) based on the wind reports only

Appendix F: WRF Model Identification of Convective Storms with Rotating Updrafts – Computation of Updraft Helicity

1. Storm Relative Environmental Helicity

Helicity, H , is a scalar measure of the potential for helical flow (i.e., the pattern of a corkscrew) to develop in a moving fluid defined by

$$H = \vec{V} \bullet \nabla \times \vec{V}.$$

Expressed in its component form,

$$H = u\left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\right) + v\left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right) + w\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right).$$

The portion of helicity associated with the *storm relative streamwise component* is that along the ambient horizontal velocity vector, or

$$H_s = -(u - c_u)\left(\frac{\partial v}{\partial z}\right) + (v - c_v)\left(\frac{\partial u}{\partial z}\right),$$

where c_u and c_v are the storm motion and terms involving w neglected. Integrating H_s vertically through the thunderstorm inflow layer, z , yields the *storm relative environmental helicity*, SREH,

$$SREH = -\int_{z_0}^z \left[(u - c_u)\left(\frac{\partial v}{\partial z}\right) - (v - c_v)\left(\frac{\partial u}{\partial z}\right) \right] dz.$$

SREH is a commonly used parameter to assess the severe thunderstorm potential of the environment and is often integrated from the surface to 1 - 3 km AGL. Order of magnitude values of SREH are $\sim O(50)$ to $O(300) \text{ m}^2/\text{s}^2$ in environments that tornadic storms.

2. Updraft Helicity

With the availability of numerical models containing sufficient resolution to resolve convective processes explicitly, it is now possible to calculate a *vertical component of helicity* associated with the convective updraft. This is the vertical integral of the third term in equation (2) and referred to as *updraft helicity*, U_H . Thus,

$$U_H = \int_{z_0}^z \left[w\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) \right] dz = \int_{z_0}^z [w\zeta] dz$$

where w is the vertical component of the relative vorticity at grid points where $w > 0$. In post processing the WRF members for the SPC/NSSL Spring Program, equation (5) is integrated vertically from $z_0 = 2$ km to $z = 5$ km AGL using a midpoint approximation. Data are available every 1000 meters AGL, so equation (5) is computed as

$$U_H = \int_{z_0}^z [w\zeta] dz \approx \sum_{z=2000m}^{z=5000m} (\overline{w\zeta} \Delta Z) = (\overline{w\zeta}_{2,3} + \overline{w\zeta}_{3,4} + \overline{w\zeta}_{4,5}) \times 1000 ,$$

where the over bar indicates a layer average and the subscripts indicate the bottom and top of the layer in kilometers. Early experience indicates that typical values of U_H associated with WRF predicted supercell thunderstorms are have U_H of at least $\sim O(50)$ m^2/s^2 and that significant supercells have $U_H \sim O(150)$ m^2/s^2 .

Appendix G: DTC Verification Metrics Summary

1. Traditional Verification Metrics – *excerpted from the WWRP/WGNE Joint Group on Forecast Verification Research website on Forecast Verification: Issues, Methods and FAQ (<http://www.cawcr.gov.au/projects/verification/>)*

1a. Statistics for dichotomous (2-category) variables

For dichotomous variables (e.g., precipitation amount above or below a threshold) on a grid, typically the forecasts are evaluated using a diagram like the one shown in Fig. 1. In this diagram, the area “**H**” represents the intersection between the forecast and observed areas, or the area of **Hits**; “**M**” represents the observed area that was missed by the forecast area, or the “**Misses**”; and “**F**” represents the part of the forecast that did not overlap an area of observed precipitation, or the “**False Alarm**” area. A fourth area is the area outside both the forecast and observed regions, which is often called the area of “**Correct Nulls**” or “**Correct Rejections**”.

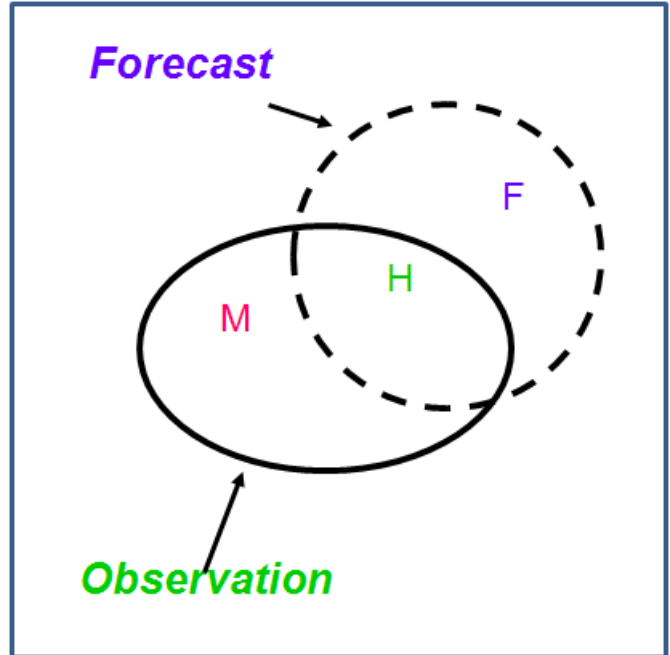


Figure 1. Diagram showing hits, misses, and false alarms for dichotomous forecast/observations.

This situation can also be represented in a “contingency table” like the one shown in Table 1. In this table the entries in each “cell” represent the counts of hit, misses, false alarms, and correct rejections. The counts in this table can be used to compute a variety of traditional verification measures, described in the following sub-sections.

Table 1. Contingency table illustrating the counts used in verification statistics for dichotomous (e.g., Yes/No) forecasts and observations. The values in parentheses illustrate the combination of forecast value (first digit) and observed value. For example, YN signifies a Yes forecast and a No observation.

Forecast	Observed		
	Yes	No	
Yes	Hits (YY)	False alarms (YN)	YY + YN
No	Misses (NY)	Correct rejections (NN)	NY + NN
	YY + NY	YN + NN	Total = YY + YN + NY + NN

Base rate

$$\text{Base rate} = \frac{\text{Hits} + \text{Misses}}{\text{Total}} = \frac{YY + NY}{\text{Total}}$$

Also known as **sample climatology** or **observed relative frequency of the event**.

Answers the question: What is the relative frequency of occurrence of the Yes event?

Range: 0 to 1.

Characteristics: Only depends on the observations. For convective weather can give an indication of how “active” a day is.

Probability of detection (POD)

$$\text{POD} = \frac{\text{Hits}}{\text{Hits} + \text{Misses}} = \frac{YY}{YY + NY}$$

Also known as **Hit Rate**.

Answers the question: What fraction of the observed Yes events was correctly forecasted?

Range: 0 to 1. **Perfect score:** 1.

Characteristics: Sensitive to hits, but ignores false alarms. Good for rare events. Can be artificially improved by issuing more Yes forecasts to increase the number of hits. Should be used in conjunction with the false alarm ratio (below) or at least one other dichotomous verification measure. POD also is an important component of the Relative Operating Characteristic (ROC) used widely for evaluation of probabilistic forecasts.

False alarm ratio (FAR)

$$\text{FAR} = \frac{\text{False alarms}}{\text{Hits} + \text{False alarms}} = \frac{YN}{YY + YN}$$

*Answers the question: What fraction of the predicted "yes" events did **not** occur (i.e., were false alarms)?*

Range: 0 to 1. **Perfect score:** 0.

Characteristics: Sensitive to false alarms, but ignores misses. Very sensitive to the climatological frequency of the event. Should be used in conjunction with the probability of

detection (above). Relative Operating Characteristic (ROC) used widely for evaluation of probabilistic forecasts.

Bias

$$\text{Bias} = \frac{\text{Hits} + \text{False alarms}}{\text{Hits} + \text{Misses}} = \frac{YY + YN}{YY + NY}$$

Also known as **Frequency Bias**.

Answers the question: How similar were the frequencies of Yes forecasts and Yes observations?

Range: 0 to infinity. **Perfect score:** 1.

Characteristics: Measures the ratio of the frequency of forecast events to the frequency of observed events. Indicates whether the forecast system has a tendency to underforecast (Bias < 1) or overforecast (Bias > 1) events. Does not measure how well the forecast gridpoints correspond to the observed gridpoints, only measures overall relative frequencies. Can be difficult to interpret when number of Yes forecasts is much larger than number of Yes observations.

Critical Success Index (CSI)

Also known as **Threat Score (TS)**.

$$\text{CSI} = \text{TS} = \frac{\text{Hits}}{\text{Hits} + \text{Misses} + \text{False alarms}} = \frac{YY}{YY + NY + YN}$$

Answers the question: How well did the forecast "yes" events correspond to the observed "yes" events?

Range: 0 to 1, 0 indicates no skill. **Perfect score:** 1.

Characteristics: Measures the fraction of observed and/or forecast events that were correctly predicted. It can be thought of as the *accuracy* when correct negatives have been removed from consideration. That is, CSI is only concerned with forecasts that are important (i.e., assuming that the correct rejections are not important). Sensitive to hits, penalizes both misses and false alarms. Does not distinguish the source of forecast error. Depends on climatological frequency of events (poorer scores for rarer events) since some hits can occur purely due to random chance. Non-linear function of POD and FAR. Should be used in combination with other contingency table statistics (e.g., Bias, POD, FAR).

Gilbert Skill Score (GSS)

Also commonly known as **Equitable Threat Score (ETS)**.

$$\text{GSS} = \text{ETS} = \frac{\text{Hits} - \text{Hits}_{\text{random}}}{\text{Hits} + \text{Misses} + \text{False alarms} - \text{Hits}_{\text{random}}} = \frac{\text{YY} - \text{YY}_{\text{random}}}{\text{YY} + \text{NY} + \text{YN} - \text{YY}_{\text{random}}}$$

where

$$\text{Hits}_{\text{random}} = \text{YY}_{\text{random}} = \frac{(\text{Hits} + \text{False alarms})(\text{Hits} + \text{Misses})}{\text{Total}} = \frac{(\text{YY} + \text{YN})(\text{YY} + \text{NY})}{\text{Total}}$$

Answers the question: *How well did the forecast "yes" events correspond to the observed "yes" events (accounting for hits that would be expected by chance)?*

Range: -1/3 to 1; 0 indicates no skill. **Perfect score:** 1.

Characteristics: Measures the fraction of observed and/or forecast events that were correctly predicted, adjusted for the frequency of hits that would be expected to occur simply by random chance (for example, it is easier to correctly forecast rain occurrence in a wet climate than in a dry climate). The GSS (ETS) is often used in the verification of rainfall in NWP models because its "equitability" allows scores to be compared more fairly across different regimes; however it is not truly equitable. Sensitive to hits. Because it penalizes both misses and false alarms in the same way, it does not distinguish the source of forecast error. Should be used in combination with at least one other contingency table statistic (e.g., Bias).

1b. Statistics for continuous forecasts and observations – excerpted from the WWRP/WGNE Joint Group on Forecast Verification Research website on Forecast Verification: Issues, Methods and FAQ (<http://www.cawcr.gov.au/projects/verification/>)

For this category of statistical measures, the grids of forecast and observed values – such as precipitation or reflectivity – are overlain on each other, and error values are computed. The grid of error values is summarized by accumulating values at all of the grid points and used to compute measures such as mean error and root mean squared error. This section is included for completeness but DTC will not be providing any continuous stats for this Spring Experiment.

These statistics are defined in the sub-sections below. In the equations in these sections, f_i signifies the **forecast value** at gridpoint i , o_i represents the observed value at gridpoint i , and N is the **total number** of gridpoints.

Mean error (ME)

$$ME = \frac{1}{N} \sum_{i=1}^N (f_i - o_i)$$

Also called the **(additive) Bias**.

Answers the question: What is the average forecast error?

Range: minus infinity to infinity. **Perfect score:** 0.

Characteristics: Simple, familiar. Measures *systematic* error. Does not measure the magnitude of the errors. Does not measure the correspondence between forecasts and observations; it is possible to get a **perfect ME score for a bad** forecast if there are compensating errors.

Pearson Correlation Coefficient (r)

$$r = \frac{\sum_{i=1}^n (f_i - \bar{f})(o_i - \bar{o})}{\sqrt{\sum_{i=1}^n (f_i - \bar{f})^2} \sqrt{\sum_{i=1}^n (o_i - \bar{o})^2}}$$

where \bar{f} is the average forecast value and \bar{o} is the average observed value.

Also called the **linear correlation coefficient**.

Answers the question: What is the linear association between the forecasts and observations?

Range: -1 to 1. **Perfect score:** 1

Characteristics: r can range between -1 and 1; a value of 1 indicates perfect correlation and a value of -1 indicates perfect negative correlation. A value of 0 indicates that the forecasts and

observations are not correlated. The correlation does not take into account the mean error, or additive bias; it only considers linear association.

Mean squared error (MSE) and root-mean squared error (RMSE)

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (f_i - o_i)^2$$

$$\text{RMSE} = \sqrt{\text{MSE}}$$

MSE can be re-written as

$$\text{MSE} = (\bar{f} - \bar{o})^2 + s_f^2 + s_o^2 - 2s_f s_o r_{fo},$$

where \bar{f} is the average forecast value, \bar{o} is the average observed value, s_f is the standard deviation of the forecast values, s_o is the standard deviation of the observed values, and r_{fo} is the correlation between the forecast and observed values. Note that $\bar{f} - \bar{o} = \text{ME}$ and $s_f^2 + s_o^2 - 2s_f s_o r_{fo}$ is the estimated variance of the error, s_{f-o}^2 . Thus, $\text{MSE} = \text{ME}^2 + s_{f-o}^2$. To understand the behavior of **MSE**, it is important to examine *both* of these terms of **MSE**, rather than examining **MSE** alone. Moreover, **MSE** can be strongly influenced by **ME**, as shown by this decomposition.

The standard deviation of the error, s_{f-o} , is simply $s_{f-o} = \sqrt{s_{f-o}^2} = \sqrt{s_f^2 + s_o^2 - 2s_f s_o r_{fo}}$.

Note that the standard deviation of the error (ESTDEV) is sometimes called the “Bias-corrected MSE” (BCMSE) because it removes the effect of overall bias from the forecast-observation squared differences.

Answers the question: *What is the average magnitude of the forecast errors?*

Range: 0 to infinity. **Perfect score:** 0.

Characteristics: Simple, familiar. Measures "average" error, weighted according to the square of the error. Does not indicate the direction of the deviations. The RMSE puts greater influence on large errors than smaller errors, which may be a good thing if large errors are especially undesirable, but may also encourage conservative forecasting.

2. MODE Summary Metrics

The Method for Object-based Diagnostic Evaluation (MODE) identifies and matches spatial objects in the forecast and observed fields. A convolution radius (r) and a precipitation/reflectivity threshold (t) are used to identify objects; different combinations of these parameters lead to objects with different characteristics, and can be used to evaluate forecasts as a function of threshold and scale.

In the object matching and merging¹ process, all possible pairs of forecast and observed objects are assigned a total “interest” value. This value is formulated from the weighted sum of specific interest values that are associated with differences in particular attributes between the forecast and observed objects. According to the current weighting scheme, the total interest value is large when objects are located close to each other and are about the same size, and is smaller for pairs of objects that are further apart and have different sizes. Note that users can specify other components of interest, and their relative weights, in the configuration file for running MODE, according to what is most relevant for their particular application.

Figure 2 illustrates a scenario in which three forecast objects and two observed objects have been identified in the two fields. The total interest values for all of the pairs of forecast and observed objects are shown in the associated table. In previous work an interest threshold of 0.70 has been found to be a reasonable indicator of a good match. Thus, in this case, forecast object 1 is a good match with both observed objects 1 and 2, and forecast object 3 matches well with observed object 2. Forecast object 3 does not match well with either of the observed objects, mostly because of its small size. Because both forecast objects 1 and 2 match observed object 2, and forecast object 1 also matches observed object 1, these objects form a matched “cluster” in the forecast and observed fields.

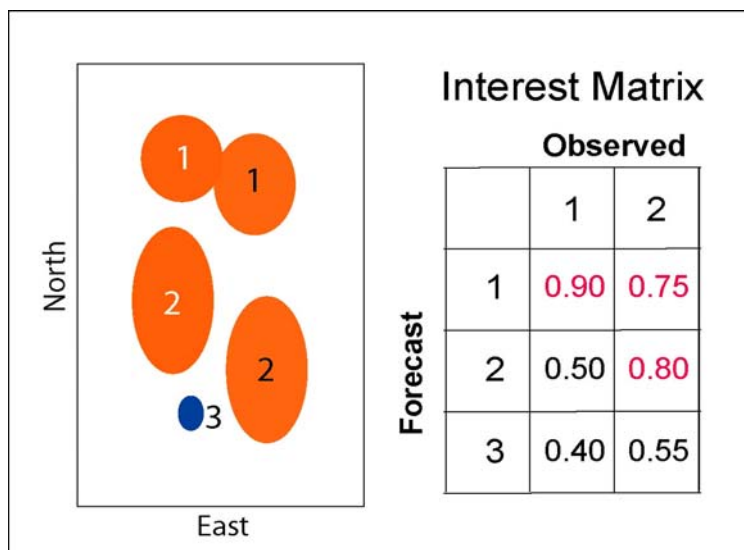


Figure 2. Schematic showing hypothetical forecast rain objects (black numerical labels) and observed rain objects (white numerical labels) with the corresponding interest matrix at right. Orange-shaded objects are matched whereas blue shading denotes no match. Total interest values greater than 0.7 are shown in red numbers in matrix. From Davis et al. (2009).

Some of the forecast attributes that are (or can be considered) in determining matches between objects include object size, distribution of intensity values, orientation angle, and location. Comparisons of these attributes, along with the total interest values, also can be used to help measure the quality of the forecast performance.

¹ “Merging” refers to the connection of objects in the same field, while “matching” refers to the connection between objects in the forecast and observed field.

Median of Maximum Interest (MMI)

This measure is computed using the total interest values for all of the pairs of objects. It considers the maximum total interest values associated with each forecast object and each observed object. From this set, the median value is computed and is the MMI.

Example: Forecast and observed objects in Fig. 2

Maximum interest values for all of the forecast and observed objects are as follows:

For forecast object 1, the maximum total interest is 0.90.

For forecast object 2, the maximum total interest is 0.80.

For forecast object 3, the maximum total interest is 0.55.

For observed object 1, the maximum total interest is 0.90.

For observed object 2, the maximum total interest is 0.80.

The median of those 5 numbers is 0.80, so $MMI = 0.80$.

This number can be small because no objects match well, or because there are many extra objects that don't match well.

Larger MMI values imply a better match between forecast and observed objects.

Area-Weighted CSI (AWCSI)

Area Weighted Critical Success Index (AWCSI)

$$AWCSI = \frac{[(hit\ area\ weight) * \#hits]}{[(hit\ area\ weight * \# hits) + (miss\ area\ weight * \# misses) + (false\ alarm\ area\ weight * \# false\ alarms)]}$$

Where each area weight is the ratio of size of the (hit, miss, or false alarm) objects to the total area of all objects and # hits = number of matched objects; # misses = # unmatched observed objects; and # false alarms = # unmatched forecast objects.

Answers the question: How well did the forecast "yes" objects correspond to the observed "yes" objects?

Range: 0 to 1, 0 indicates no skill. Perfect score: 1.

Characteristics: Measures the area-weighted fraction of observed and/or forecast events that were correctly predicted. It can be thought of as the /accuracy/ when correct negatives have been removed from consideration, that is, /TS/ is only concerned with forecasts that count. Sensitive to hits, penalizes both misses and false alarms. Does not distinguish source of forecast error.

In a grid-based CSI each gridpoint that is counted in computing the CSI contributes represents an area with the same size but with MODE objects, the various objects can have a wide variety of sizes. Thus, area weighting makes sense. and observed objects.

Median Intersection over Area (MIA)

Ratio of intersection area to union area (unitless). Ranges from zero to one: One is perfect, smaller implies less overlap. This measure is the mean for all clusters of objects with interest values greater than 0.7.

Median Area Ratio (MAR)

Ratio of the areas of two objects defined as the lesser of the forecast area divided by the observation area or its reciprocal (unitless). The ideal value is 1, since this means that the forecast and observed objects are exactly the same size. Smaller implies that the forecast was either too small or too large. This measure is the mean for all clusters of objects with interest values greater than 0.7.

Median Centroid Distance (MCD)

Distance between two objects centroids (in grid units). Smaller is better, since this means the objects are closer. This measure is the mean for all clusters of objects with interest values greater than 0.7.

Median Angle Difference (MAD)

Difference between the axis angles of two objects (in degrees). This is only meaningful if objects seem to be more linear than circular, e.g. lines of thunderstorms. When they are linear, this measure tells you how well the angle of the forecast line matches the angle of the observed line. Smaller differences are better. This measure is the mean for all clusters of objects with interest values greater than 0.7.

Intensity with confidence intervals

10th, 25th, 50th, 75th, and 90th percentiles of intensity of the filtered field within the object (various units). This tells you the distribution of values within an object (think of this as the numeric equivalent of a boxplot). There are no ideal values. However, if you compare the distribution of values within a forecast object and an observed object, you would like them to match up. We recommend checking to see how close the median and 90th percentile values are. This will tell you if you forecast is too intense or not intense enough. This measure is the mean for all clusters of objects with interest values greater than 0.7.

Median P50 Difference

First, the difference between the forecast and observed 50th percentile intensity (median) for matched objects is calculated. The median of the difference for given time is then calculated and plotted.

Median P90 Difference

First, the difference between the forecast and observed 90th percentile intensity for matched objects is calculated. The median of the difference for given time is then calculated and plotted.

Areal Coverage (ACOV)

Proportion of observation grid points inside the object. Intended to be used similar to Base Rate for traditional statistics.

3. Probabilistic Evaluation – excerpted from the WWRP/WGNE Joint Group on Forecast Verification Research website on Forecast Verification: Issues, Methods and FAQ (<http://www.cawcr.gov.au/projects/verification/>).

A probabilistic forecast gives a *probability* of an event occurring, with a value between 0 and 1 (or 0 and 100%). In general, it is difficult to verify a single probabilistic forecast. Instead, a set of probabilistic forecasts, p_i , is verified using observations that those events either occurred ($o_i=1$) or did not occur ($o_i=0$).

An accurate probability forecast system has:

- * *reliability* - agreement between forecast probability and mean observed frequency
- * *sharpness* - tendency to forecast probabilities near 0 or 1, as opposed to values clustered around the mean
- * *resolution* - ability of the forecast to resolve the set of sample events into subsets with characteristically different outcomes

Brier score -
$$BS = \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2 = \frac{1}{N} \sum_{k=1}^K n_k (p_k - \bar{o}_k)^2 - \frac{1}{N} \sum_{k=1}^K n_k (\bar{o}_k - \bar{o})^2 + \bar{o}(1 - \bar{o})$$

Brier score provides the user with a measure of the magnitude of the probability forecast errors. Measures the mean squared probability error. Murphy (1973) showed that it could be partitioned into three terms: (1) *reliability*, (2) *resolution*, and (3) *uncertainty*. These variables will also be made available during this Spring Experiment.

Range: 0 to 1. **Perfect score:** 0.

Characteristics: Sensitive to climatological frequency of the event: the more rare an event, the easier it is to get a good *BS* without having any real skill. Negative orientation (smaller score better) - can "fix" by subtracting *BS* from 1.

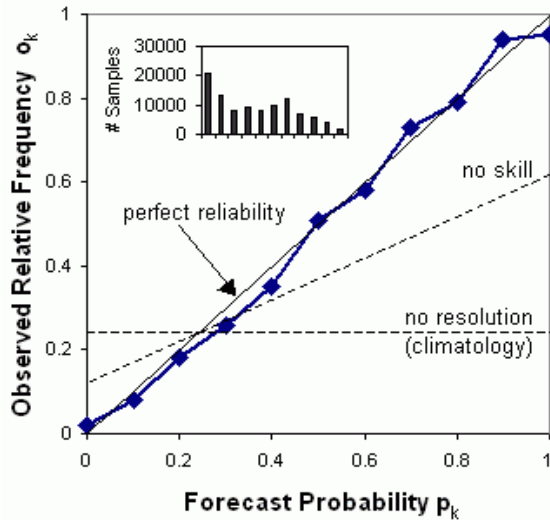
Brier skill score -
$$BSS = \frac{BS - BS_{reference}}{0 - BS_{reference}} = 1 - \frac{BS}{BS_{reference}}$$

Answers the question: What is the relative skill of the probabilistic forecast over that of climatology, in terms of predicting whether or not an event occurred?

Range: $-\infty$ to 1, 0 indicates no skill when compared to the reference forecast. **Perfect score:** 1.

Characteristics: Measures the improvement of the probabilistic forecast relative to a reference forecast (usually the long-term or sample climatology), thus taking climatological frequency into account. Not strictly proper. Unstable when applied to small data sets; the rarer the event, the larger the number of samples needed.

Reliability diagram - The reliability diagram plots the observed frequency against the forecast probability, where the range of forecast probabilities is divided into K bins (for example, 0-5%, 5-15%, 15-25%, etc.). The sample size in each bin is often included as a histogram or values beside the data points.



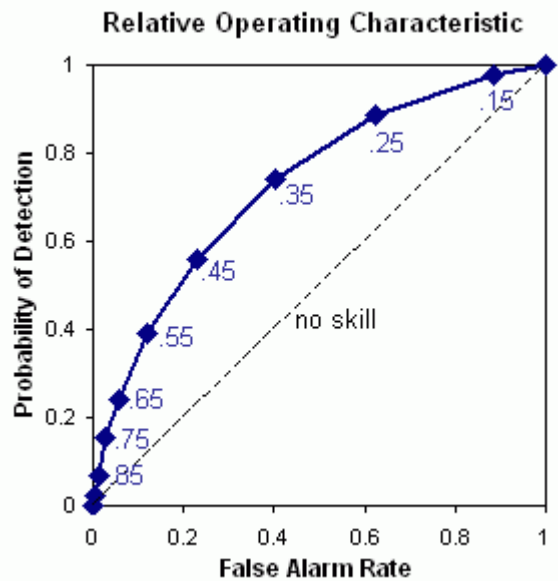
Reliability is indicated by the proximity of the plotted curve to the diagonal. The deviation from the diagonal gives the *conditional bias*. If the curve lies below the line, this indicates overforecasting (probabilities too high); points above the line indicate underforecasting (probabilities too low). The flatter the curve in the reliability diagram, the less resolution it has. A forecast of climatology does not discriminate at all between events and non-events, and thus has no resolution. Points between the "no skill" line and the diagonal contribute positively to the [Brier skill score](#). The frequency of forecasts in each probability bin (shown in the histogram) shows the sharpness of the forecast. The reliability diagram is

conditioned on the forecasts (i.e., given that X was predicted, what was the outcome?), and can be expected to give information on the real meaning of the forecast. It is a good partner to the [ROC](#), which is conditioned on the observations.

Relative operating characteristic -Plot hit rate (POD) vs false alarm rate ($POFD$), using a set of increasing probability thresholds (for example, 0.05, 0.15, 0.25, etc.) to make the yes/no decision. The area under the ROC curve is frequently used as a score.

Answers the question: What is the ability of the forecast to discriminate between events and non-events?

ROC: Perfect: Curve travels from bottom left to top left of diagram, then across to top right of diagram. Diagonal line indicates no skill.
ROC area: Range: 0 to 1, 0.5 indicates no skill.
Perfect score: 1



Characteristics: ROC measures the ability of the forecast to discriminate between two alternative outcomes, thus measuring resolution. It is not sensitive to bias in the forecast, so says

nothing about reliability. A biased forecast may still have good resolution and produce a good ROC curve, which means that it may be possible to improve the forecast through calibration. The ROC can thus be considered as a measure of potential usefulness.

The ROC is conditioned on the observations (i.e., given that Y occurred, what was the corresponding forecast?) It is therefore a good companion to the reliability diagram, which is conditioned on the forecasts. More information on ROC can be found in Mason 1982, Jolliffe and Stephenson 2003 (ch.3), and the WISE site (<http://wise.cgu.edu/stdmod/measures6.asp>).